Adrian Varallyay
October 5, 2021

QMSS-5016-Lab-1

Trend Analysis

For the trend analysis, I was interested in changes in *reported happiness* over time. In recent years there have been many arguments sighting statistical data that shows that the material conditions and the general living conditions for humans have improved vastly over the past century or so. I don't debate that; however, I question, what about the unavailable or degraded/poorly recorded or kept data? What about the human subjective experience, has that improved as well? Do those material/living condition measures correlate positively with reported happiness?

Variables

From the GSS data set, I chose ***happy*** as my main variable of interest, but also included ***age*** and ***gender*.** A quick examination of the variable ***happy*** (on the GSS data explorer website) provides the question from the survey,

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?

The responses were: 1 = Very happy, 2= Pretty happy, 3 =Not too happy. For ease of interpretability, I reverse coded ***happy*** (mean = 2.19) so that happiness increases with numeric increase. As such, *appendix A* shows cross tables and descriptive statistics for all variables starting with ***happy*** which I renamed as ***Happiness_these_days*** after reverse coding.
The distribution for respondents answering 1 (Not too happy), there were 6,510 individuals – 12.48%, for respondents answering 2 (Pretty happy), there were 29,174 individuals – 55.94%, for respondents answering 3 (Very happy), there were 16,469 individuals – 31.58%. As this is a categorical variable, I separated the response options and created dummy variables for each (***Not_too_happy***, ***Pretty_happy***, and ***Very_happy***).
I recoded ***gender*** as a dummy variable (0 = male and 1 = female) and renamed it ***female***. The distribution for males was 23,011 individuals – 44.12%, for females, there were 29,142 individuals – 55.88%.
I recoded ***age*** as a dummy variable that represented "old" age (≥50) and "not old" (≤49), and renamed it ***old***. The distribution for respondents answering "not old", there were 31,996 individuals – 61.35%, for respondents answering "old", there were 20,157 individuals – 38.65%.

Plotted Trends and OLS results

Figure 1 shows the plotted trend for ***Happiness_these_days.*** The trend for reported happiness - for years starting 1972 through 2012, seems to be declining noticeably. Interesting to note, there tends to be spikes in reported happiness around the start of a new decade, except for 2010 which shows a dip – possibly related to the temporal proximity to the stock market crash of 2008.

Adrian Varallyay
October 5, 2021

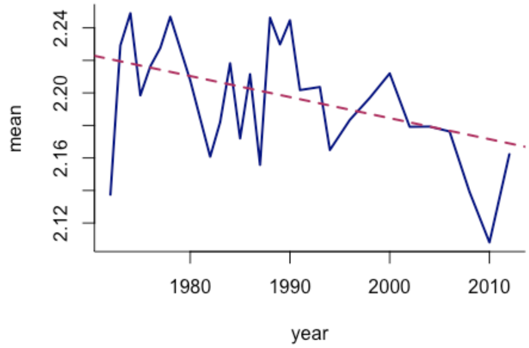

Fig 1. Trend for ***Happiness_these_days***

Appendix B shows that running a super naïve OLS on ***Happiness_these_days***, with functional form

$$Y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon_i$$

where x_i is represents ***year***, results in a very statistically significant result (p<0.001). The coefficient on year is -0.002 (t = -6.153), meaning on average, for every increase in ***year***, mean scores for happiness goes down 0.002 scale points. The adjusted R-squared is 0.0007062, meaning about 0.07% of the variation in responses is explained by the model. However, ***happy*** is a categorical variable.

Running a naïve OLS on ***Very_happy*** results in a very statistically significant result (p<0.001), with -0.001 as a coefficient on year (t = -7.151), meaning on average, for every increase in ***year,*** mean scores for happiness goes down 0.001 scale points. The adjusted R-squared is 0.0009605, meaning about 0.1% of the variation in responses is explained by the model.

Running a naïve OLS on ***Pretty_happy*** results in a very statistically significant result (p<0.001), with 0.001 as a coefficient on year (t = 5.51), meaning on average, for every increase in ***year,*** mean scores for happiness increases by 0.001 scale points. The adjusted R-squared is 0.0005627, meaning about 0.06% of the variation in responses is explained by the model.

Running a naïve OLS on ***Not _too_happy*** results in a not quite statistically significant result (p<0.1). Nonetheless, the coefficient on year is 0.0002 (t = 0.0754). The direction of the effect is similar to that of ***Pretty_happy.***

Trends with age

Figure 2a shows the plotted trend for ***Very_happy*** and figure 2b shows the same but with quadratic lines superimposed. The trend seems to maintain the noticeable decline seen in the ***Happiness_these_days*** plot. It's interesting to note the happiness disparity between old and young respondents which is maintained in the trends for both ***Very_happy*** and ***Pretty_happy,*** but not for ***Not_too_happy***.

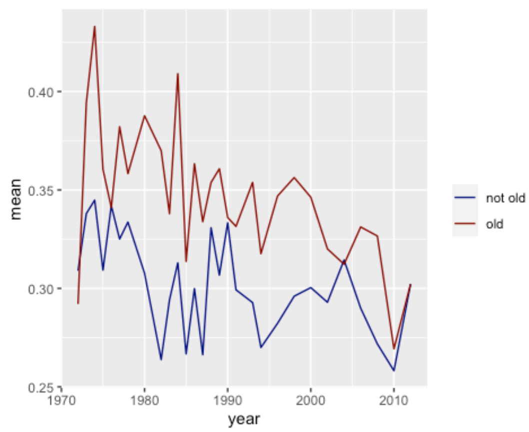

Fig 2a. Trend for ***Very_happy***

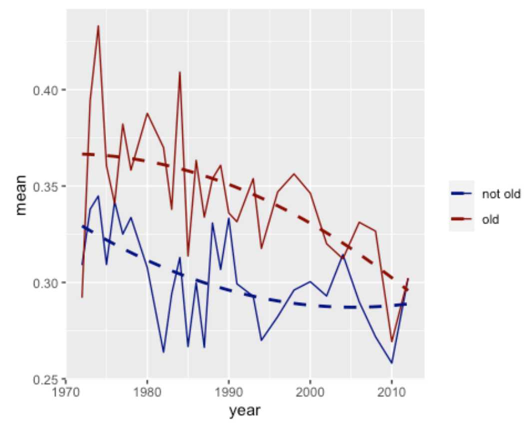

Fig 2b. Trend for ***Very_happy*** (quadratic)

Figure 3a shows the plotted trend for ***Pretty_happy*** and figure 3b shows the same but with quadratic lines superimposed. The trend does not maintain the decline seen in the ***Happiness_these_days*** plot. In fact, the trend shows a general rise in mean scores, with a start of a decline in scores at some point. It's interesting to note the disparity of reported happiness mean scores between old and young respondents is reversed for respondents answering "pretty happy", as compared to those of ***Very_happy.***

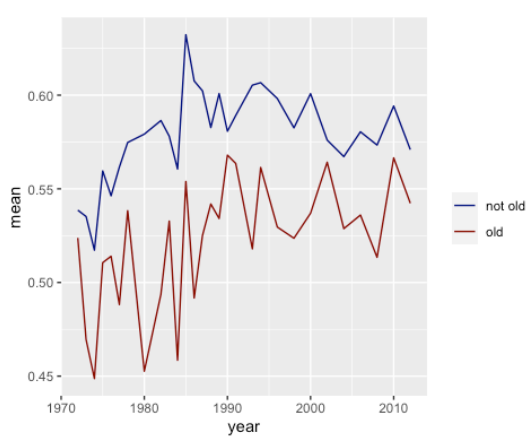

Fig 3a. Trend for ***Pretty_happy***

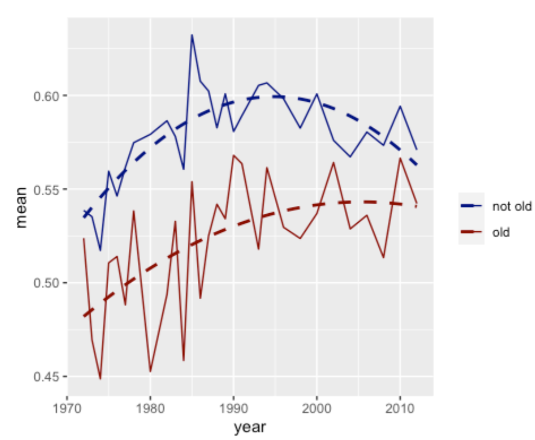

Fig 3b. Trend for ***Pretty_happy*** (quadratic)

Figure 4a shows the plotted trend for ***Not_too_happy*** and figure 4b shows the same but with quadratic lines superimposed. The trend does not maintain the decline seen in the ***Happiness_these_days*** plot. In fact, the trend differs from all previous plots. The trend for ***Not_too_happy*** shows a steep drop in mean scores followed by a steep rise at some point. It's interesting to note how much more aligned the old and young respondents are on reported happiness mean scores as compared to ***Very_happy*** and ***Pretty happy.*** The plot suggests reports of feeling "not too happy" began declining , but at some point began increasing again.

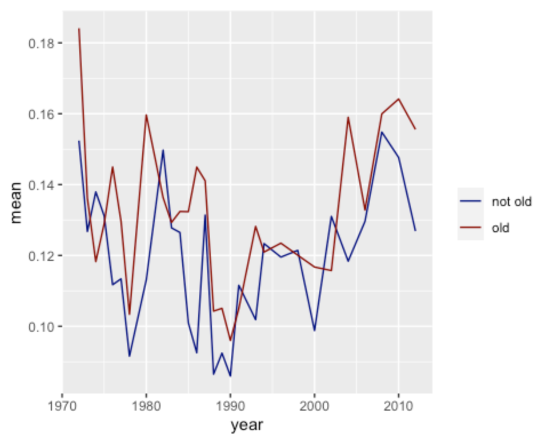

Fig 4a. Trend for ***Not_too_happy***

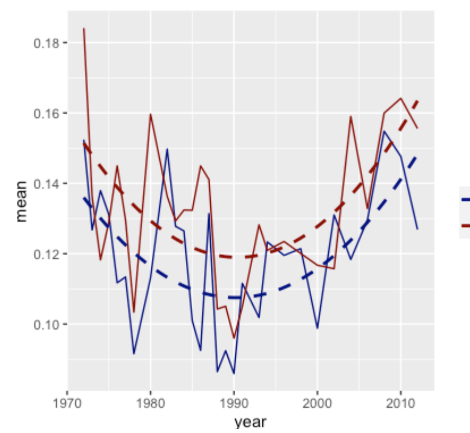

Fig 4b. Trend for ***Not_too_happy*** (quadratic)

Trends with gender

Figure 5a shows the plotted trend for ***Very_happy*** and figure 5b shows the same but with quadratic lines superimposed. The trend seems to maintain the noticeable decline seen in the ***Happiness_these_days*** plot. It's interesting to note how close, or aligned males and females are in regards to reported happiness mean scores. At first, we note a steeper drop in scores for females than males, but eventually scores converge/align better. However, towards the end of the time period, the quadratic trend line suggests a dovetail in reported happiness mean scores in the later years, with females beginning to report higher scores.

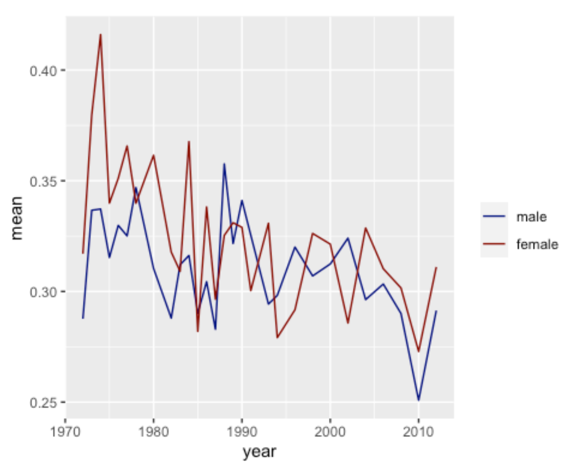

Fig 5a. Trend for ***Very_happy***

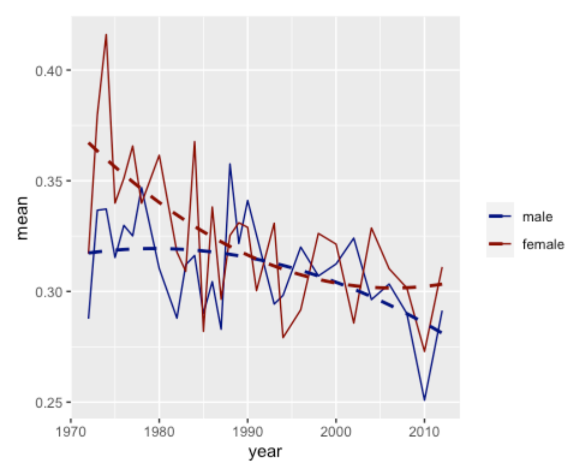

Fig 5b. Trend for ***Very_happy*** (quadratic)

Figure 6a shows the plotted trend for ***Pretty_happy*** and figure 6b shows the same but with quadratic lines superimposed. The trend does not maintain the decline seen in the ***Happiness_these_days*** plot. In fact, the trend shows the opposite, an increase in reported happiness mean scores, with a start of a decline in scores at some point. It's interesting to note again that there is not much disparity in reported happiness mean scores between male and female respondents.

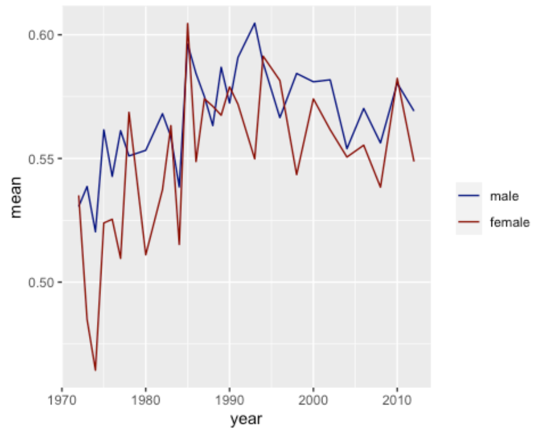

Fig 6a. Trend for ***Pretty_happy***

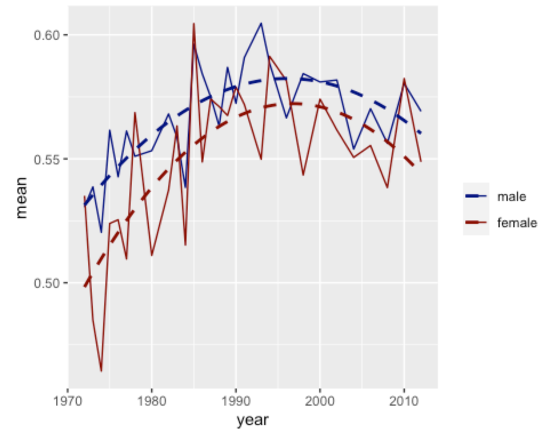

Fig 6b. Trend for ***Pretty_happy*** (quadratic)

Figure 7a shows the plotted trend for ***Not_too_happy*** and figure 7b shows the same but with quadratic lines superimposed. The trend does not maintain the decline seen in the ***Happiness_these_days*** plot. In fact, the trend differs from all previous plots. The trend for ***Not_too_happy*** shows a drop in mean scores – particularly more so for males, followed by a steep rise at some point. It's interesting to note how male and female respondents do not have too much disparity in responses across all the plots for gender. The plot suggests reports of feeling "not too happy" began declining, but at some point, began increasing again, with the effect stronger on males than females.

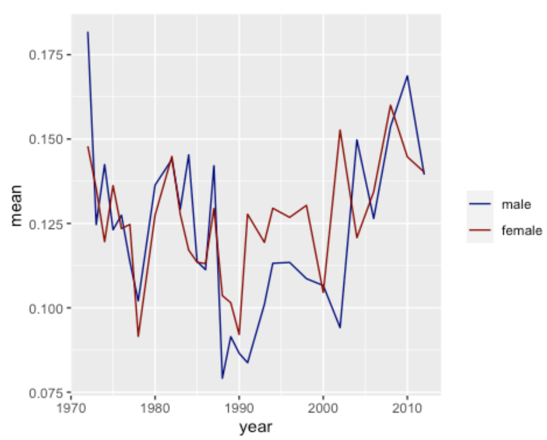

Fig 7a. Trend for ***Not_too_happy***

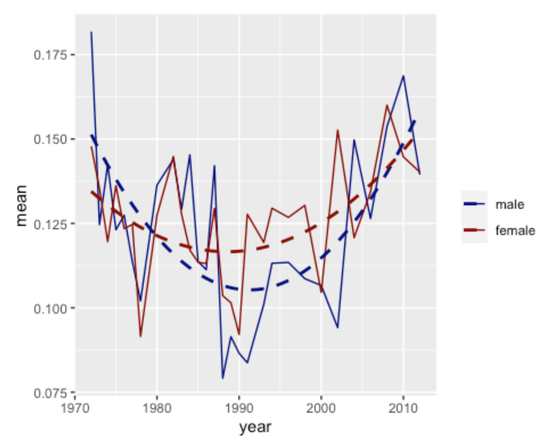

Fig 7b. Trend for ***Not_too_happy*** (quadratic)

Alternate functional forms

Appendix C shows the results of running an OLS regression on ***Very_happy***, ***Pretty_happy, and Not_too_happy,*** with functional form

Adrian Varallyay
October 5, 2021

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x^2_{1i} + \varepsilon_i$$

where x_{1i} represents ***year***, and x_{2i} represents and x_{2i} represents ***year***^{2}**.** Results for the regression on ***Very_happy*** are not statistically significant.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.411 as a coefficient on ***year*** (t = 6.178) and -0.0001 as a coefficient on ***year***^{2} (t = -6.163). The results suggest, on average, for every increase in **year,** mean scores for reported happiness increases by 0.411 scale points and then at some point each increase in ***year***^{2} produces 0.0001 decrease in mean scores for reported happiness. The adjusted R-squared is 0.001271, meaning about 0.13% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.351 as a coefficient on ***year*** (t = -7.925) and 0.00009 as a coefficient on ***year***^{2} (t = *7.930*). The results suggest, on average, for every increase in **year,** mean scores for reported happiness decreases by 0.351 scale points and then at some point each increase in ***year***^{2} produces 0.00009 increase in mean scores for reported happiness. This suggests individuals began reporting feeling less "not too happy" and then at some point began reporting feeling more "not too happy". The adjusted R-squared is 0.001227, meaning about 0.12% of the variation in responses is explained by the model.

Individual subsetted regressions

Appendix D shows the results of running Individual subsetted OLS regression on ***Very_happy***, ***Pretty_happy, and Not_too_happy.*** I ran two sets of models subsetting on ***age*** and ***gender.***

Age:Old

Results for the regression on ***Very_happy*** are not statistically significant.

Results for the regression on ***Pretty_happy*** show a statistically significant result ($p<0.05$), with 0.211 as a coefficient on ***year*** (t = 1.99) and -0.0005 as a coefficient on ***year***^{2} (t = -1.977). The results suggest, on average, for every increase in **year,** for *old* individuals as compared to young, mean scores for reported happiness increases by 0.211 scale points and then at some point each increase in ***year***^{2} produces 0.0005 decrease in mean scores for reported happiness. Thus, for *old* individuals, scores increase at first and at some point, begin to decrease. The adjusted R-squared is 0.001271, meaning about 0.13% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.385 as a coefficient on ***year*** (t = -5.349) and 0.0001 as a coefficient on ***year***^{2} (t = 5.352). The results suggest, on average, for every increase in **year,** for *old* individuals as compared to *young*, mean scores for reported happiness decreases by 0.385 scale points and then at some point each increase in ***year***^{2} produces 0.0001 increase in mean scores for reported happiness. This suggests old individuals began reporting feeling less "not too happy" and then at some point began reporting feeling more “not too happy”. The adjusted R-squared is 0.001227, meaning about 0.12% of the variation in responses is explained by the model.

Age: Not Old

Results for the regression on ***Very_happy*** show a statistically significant result ($p<0.05$), with -0.172 as a coefficient on year ($t = -2.172$) and 0.0004 as a coefficient on ***year²*** ($t = 2.159$). The results suggest, on average, for every increase in **year,** for *young* individuals as compared to *old*, mean scores for reported happiness decreases by 0.172 scale points and then at some point each increase in ***year²*** produces 0.0004 increase in mean scores for reported happiness. This suggests *young* individuals began reporting feeling less “very happy” and then at some point began reporting feeling more “very happy”. The adjusted R-squared is 0.0007692, meaning about 0.08% of the variation in responses is explained by the model.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.487 as a coefficient on ***year*** ($t = 5.702$) and -0.0001 as a coefficient on ***year²*** ($t = -5.692$). The results suggest, on average, for every increase in **year,** for *young* individuals as compared to *old*, mean scores for reported happiness increases by 0.487 scale points and then at some point each increase in ***year²*** produces 0.0001 decrease in mean scores for reported happiness. Thus, for *young* individuals, scores increase at first and at some point, begin to decrease. The adjusted R-squared is 0.001344, meaning about 0.13% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.315 as a coefficient on ***year*** ($t = -5.598$) and 0.00008 as a coefficient on ***year²*** ($t = 5.602$). The results suggest, on average, for every increase in **year,** for *young* individuals as compared to *old*, mean scores for reported happiness decreases by 0.315 scale points and then at some point each increase in ***year²*** produces 0.00008 increase in mean scores for reported happiness. This suggests *young* individuals began reporting feeling less “not too happy” and then at some point began reporting feeling more “not too happy”. The adjusted R-squared is 0.0009623, meaning about 0.1% of the variation in responses is explained by the model.

Gender:Female

Results for the regression on ***Very_happy*** show a statistically significant result ($p<0.01$), with -0.240 as a coefficient on year ($t = -2.849$) and 0.0006 as a coefficient on ***year²*** ($t = 2.830$). The results suggest, on average, for every increase in **year,** for *females* as compared to *males*, mean scores for reported happiness decreases by 0.240 scale points and then at some point each increase in ***year²*** produces 0.0006 increase in mean scores for reported happiness. This suggests *females* began reporting feeling less “very happy” and then at some point began reporting feeling more “very happy”. The adjusted R-squared is 0.00177, meaning about 0.18% of the variation in responses is explained by the model.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.479 as a coefficient on ***year*** ($t = 5.326$) and -0.0001 as a coefficient on ***year²*** ($t = -5.312$). The results suggest, on average, for every increase in **year,** for *females* as compared to *males*, mean scores for reported happiness increases by 0.479 scale points and then at some point each increase in ***year²*** produces 0.0001 decrease in mean scores for reported happiness. Thus, for *females*, scores increase at first and at some point, begin to decrease. The adjusted R-squared is 0.001697, meaning about 0.17% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.238 as a coefficient on ***year*** ($t = -3.966$) and 0.00006 as a coefficient on ***year²*** ($t = 3.973$). The results suggest, on average, for every increase in **year,** for *females* as compared to *males*, mean scores for reported happiness decreases by 0.238 scale points and then at some point each increase in ***year²*** produces 0.00006 increase in mean scores for reported happiness. This suggests *females* began reporting feeling less "not too happy" and then at some point began reporting feeling more "not too happy". The adjusted R-squared is 0.000649, meaning about 0.06% of the variation in responses is explained by the model.

Gender:Male

Results for the regression on ***Very_happy*** are not statistically significant.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.343 as a coefficient on ***year*** ($t = 3.467$) and -0.00009 as a coefficient on ***year²*** ($t = -3.459$). The results suggest, on average, for every increase in **year,** for *males* as compared to *females*, mean scores for reported happiness increases by 0.343 scale points and then at some point each increase in ***year²*** produces 0.00009 decrease in mean scores for reported happiness. Thus, for *males*, scores increase at first and at some point, begin to decrease. The adjusted R-squared is 0.0007973, meaning about 0.08% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.489 as a coefficient on ***year*** ($t = -7.459$) and 0.0001 as a coefficient on ***year²*** ($t = 7.459$). The results suggest, on average, for every increase in **year,** for *males* as compared to *females*, mean scores for reported happiness decreases by 0.489 scale points and then at some point each increase in ***year²*** produces 0.0001 increase in mean scores for reported happiness. This suggests *males* began reporting feeling less "not too happy" and then at some point began reporting feeling more "not too happy". The adjusted R-squared is 0.002326, meaning about 0.23% of the variation in responses is explained by the model.

Controls

Appendix E shows the results of running running an OLS regression on ***Very_happy***, ***Pretty_happy, and Not_too_happy*** controlling for age or gender.

Gender

Results for the regression on ***Very_happy*** show a statistically significant result ($p<0.01$), with 0.012 as a coefficient on ***female*** ($t = 2.999$). The results suggest, on average, for *females* as compared to *males*, mean scores for reported happiness increases by 0.012 scale points. The adjusted R-squared is 0.001112, meaning about 0.11% of the variation in responses is explained by the model.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.416 as a coefficient on ***year*** ($t = 6.257$), -0.0001 as a coefficient on ***year²*** ($t = -6.242$), and -0.016 as a coefficient on ***female***. The results suggest, on average, for every increase in **year,** mean scores for reported happiness increases by 0.416 scale points and then at some point each increase in ***year²*** produces 0.0001 decrease in mean scores for reported happiness. Thus, scores increase at first and at some point, begin to decrease. Moreover, on average, females as compared to males, report 0.16 scale points lower on reported happiness. The adjusted R-squared is 0.001505, meaning about 0.15% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.352 as a coefficient on ***year*** ($t = -7.950$), 0.00009 as a coefficient on ***year²*** ($t = 7.955$), and 0.004 as a coefficient on ***female.*** The results suggest, on average, for every increase in **year,** mean scores for reported happiness decreases by 0.352 scale points and then at some point each increase in ***year²*** produces 0.00009 increase in mean scores for reported happiness. Thus, scores decrease at first and at some point, begin to increase. Moreover, on average, females as compared to males, report 0.004 scale points lower on reported happiness. The adjusted R-squared is 0.001237, meaning about 0.12% of the variation in responses is explained by the model.

Age

Results for the regression on ***Very_happy*** show a very statistically significant result ($p<0.001$), with 0.045 as a coefficient on ***old*** ($t = 10.699$). The results suggest, on average, for *old* individulas as compared to *young*, mean scores for reported happiness increases by 0.045 scale points. The adjusted R-squared is 0.003127, meaning about 0.31% of the variation in responses is explained by the model.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.372 as a coefficient on ***year*** ($t = 5.594$), -0.00009 as a coefficient on ***year²*** ($t = -5.577$), and -0.056 as a coefficient on ***old*** ($t = -12.652$). The results suggest, on average, for every increase in **year,** mean scores for reported happiness increases by 0.372 scale points and then at some point each increase in ***year²*** produces 0.0009 decrease in mean scores for reported happiness. Thus, scores increase at first and at some point, begin to decrease. Moreover, on average, *old* individuals as compared to *young*, report 0.056 scale points lower on reported happiness. The adjusted R-squared is 0.004308, meaning about 0.43% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a very statistically significant result ($p<0.001$), with -0.342 as a coefficient on ***year*** ($t = -7.734$), 0.00009 as a coefficient on ***year²*** ($t = 7.739$), and 0.012 as a coefficient on ***old.*** The results suggest, on average, for every increase in **year,** mean scores for reported happiness decreases by 0.342 scale points and then at some point each increase in ***year²*** produces 0.00009 increase in mean scores for reported happiness.

Thus, scores decrease at first and at some point, begin to increase. Moreover, on average, *old* individuals as compared to *young*, report 0.012 scale points lower on reported happiness. The adjusted R-squared is 0.001505, meaning about 0.15% of the variation in responses is explained by the model.

Interacted models

Moving forward, I will only focus on interaction models including ***gender*** with functional form

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x^2_{1i} + \beta_3 x_{2i} + \beta_1 x_{1i} * \beta_3 x_{2i} + \beta_2 x^2_{1i} * \beta_3 x_{2i} + \varepsilon_i$$

where $\beta_3 x_{2i}$ represents ***female***, and $\beta_1 x_{1i} * \beta_3 x_{2i}$ represents the interaction of gender and year and $\beta_2 x^2_{1i} * \beta_3 x_{2i}$ represents the interaction of gender and ***year²***.

Appendix F shows the statistically significant result ($p<0.01$) for the regression on ***Very_happy***, where 385.617 is the coefficient on ***female*** (t = 3.092), -0.386 as a coefficient on ***year:female*** interaction term (t = -3.086), and 0.00009 ***year²:female*** interaction term (t = 3.080). While females start higher with reported happiness (in year zero), on average reported happiness decreases by -0.24 scale points (male: 0.14) and at some point, increases by 0.00006 scale points (male: -0.00004). The adjusted R-squared is 0.001349, meaning about 0.13% of the variation in responses is explained by the model.

Results for the regression on ***Pretty_happy*** show a very statistically significant result ($p<0.001$), with 0.343 as a coefficient on ***year*** (t = 3.461), and -0.0001 as a coefficient on ***year²*** (t = -6.242). The results suggest, on average, for every increase in **year,** mean scores for reported happiness increases by 0.343 scale points and then at some point each increase in ***year²*** produces 0.00009 decrease in mean scores for reported happiness. Thus, scores increase at first and at some point begin to decrease. The adjusted R-squared is 0.001514, meaning about 0.15% of the variation in responses is explained by the model.

Results for the regression on ***Not _too_happy*** show a statistically significant result ($p<0.001$), with -0.489 as a coefficient on ***year*** (t = -7.410), 0.0001 as a coefficient on ***year²*** (t = 7.411). Other variables had statistically significant result ($p<0.01$) with -250.48 as a coefficient on ***female*** (t = -2.825), 0.251 as a coefficient on ***year:female*** interaction term (t = 2.821), and -0.00006 as a coefficient ***year²:female*** interaction term (t = -2.817). While females start lower with reported happiness (in year zero), on average, reported happiness decreases by -0.24 scale points (male: -0.489) and at some point, increases by 0.00006 scale points (male: -0.00012). The adjusted R-squared is 0.001383, meaning about 0.14% of the variation in responses is explained by the model.

ANOVA

Appendix G shows the results of the analysis of variance (ANOVA) of the three sets of models using ***gender.***

Very_happy

Adrian Varallyay
October 5, 2021

For Model 1, I regressed ***year***, ***year^2*, *and*** female on ***Very_happy,*** and for Model 2, I regressed ***year***, ***year^2*,** female***, year:female,*** and ***year^2:female*** on ***Very_happy***. The results are very statistically significant ($p<0.001$) with an F-statistic of 7.2097*,* which means I reject the null, Model 2 is preferred.

Pretty_happy

For Model 1, I regressed ***year***, ***year^2*, *and*** female on ***Pretty_happy,*** and for Model 2, I regressed ***year***, ***year^2*,** female***, year:female,*** and ***year^2:female*** on ***Pretty_happy***. The results are not statistically significant, I do not reject the null.

Not_too_happy

For Model 1, I regressed ***year***, ***year^2*, *and*** female on ***Pretty_happy,*** and for Model 2, I regressed ***year***, ***year^2*,** female***, year:female,*** and ***year^2:female*** on ***Pretty_happy***. The results are statistically significant ($p<0.01$) with an F-statistic of 4.8127*,* which means I reject the null, Model 2 is preferred.

Alternative variables to examine

As I considered variables to examine, I realized that there are an additional 12 I would have liked to evaluate. ***Attend*** asks, "How often do you attend religious services?". ***Degree*** inquires about respondent's academic degree. ***Marital*** asks, "Are you currently--married, widowed, divorced, separated, or have you never been married?". ***Partyid*** asks, "Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?". ***Race*** inquires about respondent's race. ***Realinc*** inquires about respondent's family income. R***elig*** asks, "What is your religious preference? Is it Protestant, Catholic, Jewish, some other religion, or no religion?". Associated with that last variable is ***reliten*** which asks, "Would you call yourself a strong (PREFERENCE NAMED IN RELIG) or a not very strong (PREFERENCE NAMED IN RELIG)?". ***Spwrkslf*** asks if respondent's spouse is self-employed or works for somebody. ***Spwrksta*** inquires about Spouse labor force status. ***Wrkslf asks*** if respondent is self-employed or works for somebody. And finally, ***Wrkstat*** asks, "Last week were you working full time, part time, going to school, keeping house, or what?". All these variables seem like they would play a strong role in how someone would respond to the ***happy*** variable, or in our case, the ***Very_happy, Pretty_happy,*** and ***Not_too_happy*** dumby variables.

Adrian Varallyay
October 5, 2021

Appendix A

Happiness_these_days

```
##    Cell Contents
## |-------------------------|
## |                   Count |
## |             Row Percent |
## |-------------------------|
##
## Total Observations in Table:  52153
##
##           |         1  |         2  |         3  |
##           |-----------|-----------|-----------|
##           |      6510  |     29174  |     16469  |
##           |   12.483% |   55.939% |   31.578% |
##           |-----------|-----------|-----------|
##
##
##    vars     n mean   sd median trimmed mad min max range  skew kurtosis se
## X1    1 52153 2.19 0.64      2    2.24   0   1   3     2 -0.18    -0.63  0
```

Very_happy

```
##    Cell Contents
## |-------------------------|
## |                   Count |
## |             Row Percent |
## |-------------------------|
##
## Total Observations in Table:  52153
##
##           |     FALSE  |      TRUE  |
##           |-----------|-----------|
##           |     35684  |     16469  |
##           |   68.422% |   31.578% |
##           |-----------|-----------|
##
##    vars     n mean   sd median trimmed mad min max range skew kurtosis se
## X1    1 52153 0.32 0.46      0    0.27   0   0   1     1 0.79    -1.37  0
```

Pretty_happy

```
##    Cell Contents
## |-------------------------|
## |                   Count |
## |             Row Percent |
## |-------------------------|
##
## Total Observations in Table:  52153
##
##           |     FALSE  |      TRUE  |
##           |-----------|-----------|
##           |     22979  |     29174  |
##           |   44.061% |   55.939% |
##           |-----------|-----------|
##
##
##    vars     n mean  sd median trimmed mad min max range  skew kurtosis se
## X1    1 52153 0.56 0.5      1    0.57   0   0   1     1 -0.24    -1.94  0
```

Adrian Varallyay
October 5, 2021

Not_too_happy female

```
##    Cell Contents
## |-------------------------|
## |                   Count |
## |             Row Percent |
## |-------------------------|
##
## Total Observations in Table:  52153
##
##           |    FALSE  |     TRUE  |
##           |-----------|-----------|
##           |     45643 |      6510 |
##           |   87.517% |   12.483% |
##           |-----------|-----------|
##
##
##    vars     n mean   sd median trimmed mad min max range skew kurtosis se
## X1    1 52153 0.12 0.33      0    0.03   0   0   1     1 2.27     3.15  0
```

Female

```
##    Cell Contents
## |-------------------------|
## |                   Count |
## |             Row Percent |
## |-------------------------|
##
## Total Observations in Table:  52153
##
##           |         0 |         1 |
##           |-----------|-----------|
##           |     23011 |     29142 |
##           |   44.122% |   55.878% |
##           |-----------|-----------|
##
##
##    vars     n mean  sd median trimmed mad min max range  skew kurtosis se
## X1    1 52153 0.56 0.5      1    0.57   0   0   1     1 -0.24    -1.94  0
```

Old

```
##    Cell Contents
## |-------------------------|
## |                   Count |
## |             Row Percent |
## |-------------------------|
##
## Total Observations in Table:  52153
##
##           |         0 |         1 |
##           |-----------|-----------|
##           |     31996 |     20157 |
##           |   61.350% |   38.650% |
##           |-----------|-----------|
##
##
##    vars     n mean   sd median trimmed mad min max range skew kurtosis se
## X1    1 52153 0.39 0.49      0    0.36   0   0   1     1 0.47    -1.78  0
```

Adrian Varallyay
October 5, 2021

Appendix B

Fig 1. Happiness_these_days

HTD

```
## Call:
## lm(formula = happiness_these_days ~ year, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2190 -0.2043 -0.1808  0.7898  0.8397
##
## Coefficients:
##                Estimate Std. Error t value             Pr(>|t|)
## (Intercept)  5.1133769  0.4749743  10.766 < 0.0000000000000002 ***
## year        -0.0014677  0.0002385  -6.153       0.000000000766 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6355 on 52151 degrees of freedom
## Multiple R-squared:  0.0007254,  Adjusted R-squared:  0.0007062
## F-statistic: 37.86 on 1 and 52151 DF,  p-value: 0.0000000007663
```

VH

```
## Call:
## lm(formula = very_happy ~ year, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3396 -0.3221 -0.3047  0.6679  0.7103
##
## Coefficients:
##                Estimate Std. Error t value             Pr(>|t|)
## (Intercept)  2.7989739  0.3472471   8.060 0.000000000000000776 ***
## year        -0.0012471  0.0001744  -7.151 0.000000000000871477 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4646 on 52151 degrees of freedom
## Multiple R-squared:  0.0009796,  Adjusted R-squared:  0.0009605
## F-statistic: 51.14 on 1 and 52151 DF,  p-value: 0.0000000000008715
```

PH

```
## Call:
## lm(formula = pretty_happy ~ year, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5808 -0.5562  0.4233  0.4428  0.4602
##
## Coefficients:
##                Estimate Std. Error t value    Pr(>|t|)
## (Intercept) -1.4845709  0.3709524  -4.002 0.000062882 ***
## year         0.0010266  0.0001863   5.510 0.000000036 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Adrian Varallyay
October 5, 2021

```
##
## Residual standard error: 0.4963 on 52151 degrees of freedom
## Multiple R-squared:  0.0005818,  Adjusted R-squared:  0.0005627
## F-statistic: 30.36 on 1 and 52151 DF,  p-value: 0.00000003602
```

NTH

```
## Call:
## lm(formula = not_too_happy ~ year, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1294 -0.1263 -0.1241 -0.1217  0.8794
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.3144030  0.2470276  -1.273   0.2031
## year         0.0002206  0.0001241   1.778   0.0754 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3305 on 52151 degrees of freedom
## Multiple R-squared:  6.062e-05,  Adjusted R-squared:  4.145e-05
## F-statistic: 3.162 on 1 and 52151 DF,  p-value: 0.0754
```

Adrian Varallyay
October 5, 2021

Appendix C

Alternate functional forms

VH

```
## Call:
## lm(formula = very_happy ~ year + yearsq, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3433 -0.3206 -0.3037  0.6672  0.7062
##
## Coefficients:
##                Estimate  Std. Error t value Pr(>|t|)
## (Intercept) 61.31883327 62.03172961   0.989    0.323
## year        -0.06001548  0.06229438  -0.963    0.335
## yearsq       0.00001475  0.00001564   0.943    0.345
##
## Residual standard error: 0.4646 on 52150 degrees of freedom
## Multiple R-squared:  0.0009967,  Adjusted R-squared:  0.0009584
## F-statistic: 26.01 on 2 and 52150 DF,  p-value: 0.0000000000051
```

PH

```
## Call:
## lm(formula = pretty_happy ~ year + yearsq, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5765 -0.5627  0.4237  0.4369  0.4858
##
## Coefficients:
##                 Estimate   Std. Error t value       Pr(>|t|)
## (Intercept) -409.7250167   66.2428677  -6.185 0.000000000625 ***
## year           0.4110004    0.0665233   6.178 0.000000000653 ***
## yearsq        -0.0001029    0.0000167  -6.163 0.000000000720 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4961 on 52150 degrees of freedom
## Multiple R-squared:  0.001309,   Adjusted R-squared:  0.001271
## F-statistic: 34.18 on 2 and 52150 DF,  p-value: 0.000000000000001461
```

NTH

```
## Call:
## lm(formula = not_too_happy ~ year + yearsq, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1541 -0.1310 -0.1179 -0.1134  0.8872
##
## Coefficients:
##                 Estimate   Std. Error t value            Pr(>|t|)
## (Intercept) 349.40618346  44.10245757   7.923 0.00000000000000237 ***
## year         -0.35098488   0.04428919  -7.925 0.00000000000000233 ***
## yearsq        0.00008817   0.00001112   7.930 0.00000000000000224 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.3303 on 52150 degrees of freedom
## Multiple R-squared:  0.001265,   Adjusted R-squared:  0.001227
## F-statistic: 33.02 on 2 and 52150 DF, p-value: 0.000000000000004645
```

Adrian Varallyay
October 5, 2021

Appendix D

Individual subsetted regressions: old

```
## Call:
## lm(formula = very_happy ~ year + yearsq, data = gss, subset = old == 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3640 -0.3555 -0.3272  0.6395  0.7027
##
## Coefficients:
##                 Estimate    Std. Error t value Pr(>|t|)
## (Intercept) -171.67225119  100.37937359  -1.710   0.0872 .
## year           0.17437780    0.10079028   1.730   0.0836 .
## yearsq        -0.00004419    0.00002530  -1.747   0.0807 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4742 on 20154 degrees of freedom
## Multiple R-squared:  0.001906,   Adjusted R-squared:  0.001807
## F-statistic: 19.24 on 2 and 20154 DF,  p-value: 0.00000000447
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq, data = gss, subset = old == 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5423 -0.5281  0.4577  0.4687  0.5166
##
## Coefficients:
##                 Estimate    Std. Error t value Pr(>|t|)
## (Intercept) -211.12480966  105.64597622  -1.998   0.0457 *
## year           0.21109097    0.10607844   1.990   0.0466 *
## yearsq        -0.00005263    0.00002663  -1.977   0.0481 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4991 on 20154 degrees of freedom
## Multiple R-squared:  0.00137,    Adjusted R-squared:  0.001271
## F-statistic: 13.82 on 2 and 20154 DF,  p-value: 0.000001001
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq, data = gss, subset = old == 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1626 -0.1414 -0.1245 -0.1195  0.8812
##
## Coefficients:
##                Estimate   Std. Error t value     Pr(>|t|)
## (Intercept) 383.79706086  71.77604496   5.347 0.0000000903 ***
## year         -0.38546877   0.07206986  -5.349 0.0000000896 ***
## yearsq        0.00009682   0.00001809   5.352 0.0000000881 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3391 on 20154 degrees of freedom
## Multiple R-squared:  0.001485,	Adjusted R-squared:  0.001386
## F-statistic: 14.99 on 2 and 20154 DF,  p-value: 0.0000003138
```

Individual subsetted regressions: Not old

```
## Call:
## lm(formula = very_happy ~ year + yearsq, data = gss, subset = old == 0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3291 -0.2995 -0.2879  0.6834  0.7138
##
## Coefficients:
##                 Estimate   Std. Error t value Pr(>|t|)
## (Intercept) 172.91470220  78.98986525   2.189   0.0286 *
## year         -0.17231962   0.07933249  -2.172   0.0299 *
## yearsq        0.00004300   0.00001992   2.159   0.0309 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4576 on 31993 degrees of freedom
## Multiple R-squared:  0.0008316,	Adjusted R-squared:  0.0007692
## F-statistic: 13.31 on 2 and 31993 DF,  p-value: 0.00000166
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq, data = gss, subset = old == 0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5993 -0.5821  0.4011  0.4156  0.4647
##
## Coefficients:
##                  Estimate    Std. Error t value     Pr(>|t|)
## (Intercept) -485.56807244   85.10981171  -5.705 0.0000000117 ***
## year           0.48740704    0.08547898   5.702 0.0000000119 ***
## yearsq        -0.00012216    0.00002146  -5.692 0.0000000127 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.493 on 31993 degrees of freedom
## Multiple R-squared:  0.001407,	Adjusted R-squared:  0.001344
## F-statistic: 22.54 on 2 and 31993 DF,  p-value: 0.0000000001658
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq, data = gss, subset = old == 0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1470 -0.1254 -0.1142 -0.1100  0.8906
##
## Coefficients:
```

Adrian Varallyay
October 5, 2021

```
##                  Estimate   Std. Error t value     Pr(>|t|)
## (Intercept) 313.65337023  56.03837419   5.597 0.0000000220 ***
## year         -0.31508742   0.05628144  -5.598 0.0000000218 ***
## yearsq        0.00007916   0.00001413   5.602 0.0000000214 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3246 on 31993 degrees of freedom
## Multiple R-squared:  0.001025,   Adjusted R-squared:  0.0009623
## F-statistic: 16.41 on 2 and 31993 DF,  p-value: 0.00000007532
```

Individual subsetted regressions: female

```
## Call:
## lm(formula = very_happy ~ year + yearsq, data = gss, subset = female == 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3669 -0.3210 -0.3045  0.6609  0.6984
##
## Coefficients:
##                  Estimate   Std. Error t value Pr(>|t|)
## (Intercept) 241.29323359  84.02093660   2.872  0.00408 **
## year         -0.24038761   0.08437294  -2.849  0.00439 **
## yearsq        0.00005995   0.00002118   2.830  0.00466 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4665 on 29139 degrees of freedom
## Multiple R-squared:  0.001839,   Adjusted R-squared:  0.00177
## F-statistic: 26.84 on 2 and 29139 DF,  p-value: 0.000000000002265
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq, data = gss, subset=female ==1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5724 -0.5559  0.4277  0.4414  0.5011
##
## Coefficients:
##                   Estimate    Std. Error t value     Pr(>|t|)
## (Intercept) -477.19856895   89.47892994  -5.333 0.0000000973 ***
## year           0.47854286    0.08985380   5.326 0.0000001013 ***
## yearsq        -0.00011983    0.00002256  -5.312 0.0000001091 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4968 on 29139 degrees of freedom
## Multiple R-squared:  0.001766,   Adjusted R-squared:  0.001697
## F-statistic: 25.78 on 2 and 29139 DF,  p-value: 0.000000000006543
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq, data = gss, subset=female ==1)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -0.1508 -0.1287 -0.1211 -0.1182  0.8822
##
## Coefficients:
##                  Estimate   Std. Error t value  Pr(>|t|)
## (Intercept) 236.90533536  59.79321134   3.962 0.0000745 ***
## year         -0.23815525   0.06004371  -3.966 0.0000732 ***
## yearsq        0.00005988   0.00001507   3.973 0.0000712 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.332 on 29139 degrees of freedom
## Multiple R-squared:  0.0007176,  Adjusted R-squared:  0.000649
## F-statistic: 10.46 on 2 and 29139 DF,  p-value: 0.0000287
```

Individual subsetted regressions: male

```
## Call:
## lm(formula = very_happy ~ year + yearsq, data = gss, subset = female ==0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3184 -0.3164 -0.3041  0.6823  0.7186
##
## Coefficients:
##                   Estimate    Std. Error t value Pr(>|t|)
## (Intercept) -144.32347367   91.98111946  -1.569    0.117
## year           0.14607953    0.09237553   1.581    0.114
## yearsq        -0.00003688    0.00002319  -1.590    0.112
##
## Residual standard error: 0.462 on 23008 degrees of freedom
## Multiple R-squared:  0.0005542,  Adjusted R-squared:  0.0004673
## F-statistic: 6.379 on 2 and 23008 DF,  p-value: 0.0017
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq, data = gss, subset=female == 0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5821 -0.5701  0.4183  0.4286  0.4681
##
## Coefficients:
##                   Estimate    Std. Error t value Pr(>|t|)
## (Intercept) -342.06229068   98.58852419  -3.470 0.000522 ***
## year           0.34330217    0.09901127   3.467 0.000527 ***
## yearsq        -0.00008599    0.00002486  -3.459 0.000543 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4952 on 23008 degrees of freedom
## Multiple R-squared:  0.0008842,  Adjusted R-squared:  0.0007973
## F-statistic: 10.18 on 2 and 23008 DF,  p-value: 0.00003807
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq, data = gss, subset=female==0)
##
## Residuals:
```

Adrian Varallyay
October 5, 2021

```
##     Min      1Q  Median      3Q     Max 
## -0.1580 -0.1321 -0.1146 -0.1068  0.8940 
## 
## Coefficients:
##                 Estimate   Std. Error t value           Pr(>|t|)    
## (Intercept) 487.38576435  65.33215403   7.460 0.0000000000000895 ***
## year         -0.48938171   0.06561230  -7.459 0.0000000000000905 ***
## yearsq        0.00012287   0.00001647   7.459 0.0000000000000902 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.3281 on 23008 degrees of freedom
## Multiple R-squared:  0.002413,   Adjusted R-squared:  0.002326 
## F-statistic: 27.83 on 2 and 23008 DF,  p-value: 0.0000000000008497
```

Adrian Varallyay
October 5, 2021

Appendix E

Female control model

```
## Call:
## lm(formula = very_happy ~ year + yearsq + female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3490 -0.3205 -0.3037  0.6688  0.7128
##
## Coefficients:
##                Estimate  Std. Error t value Pr(>|t|)
## (Intercept) 65.38698086 62.04180288   1.054  0.29193
## year        -0.06410584  0.06230453  -1.029  0.30353
## yearsq       0.00001578  0.00001564   1.009  0.31304
## female       0.01229119  0.00409801   2.999  0.00271 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4646 on 52149 degrees of freedom
## Multiple R-squared:  0.001169,   Adjusted R-squared:  0.001112
## F-statistic: 20.34 on 3 and 52149 DF,  p-value: 0.0000000000003597
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq + female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5855 -0.5603  0.4157  0.4380  0.4932
##
## Coefficients:
##                 Estimate   Std. Error t value       Pr(>|t|)
## (Intercept) -414.9932486   66.2509357  -6.264 0.000000000378 ***
## year           0.4162974    0.0665315   6.257 0.000000000395 ***
## yearsq        -0.0001042    0.0000167  -6.242 0.000000000436 ***
## female        -0.0159170    0.0043760  -3.637       0.000276 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4961 on 52149 degrees of freedom
## Multiple R-squared:  0.001563,   Adjusted R-squared:  0.001505
## F-statistic:  27.2 on 3 and 52149 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq + female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1557 -0.1316 -0.1179 -0.1135  0.8893
##
## Coefficients:
##                 Estimate   Std. Error t value            Pr(>|t|)
## (Intercept) 350.60626770  44.11276875   7.948 0.00000000000000193 ***
## year         -0.35219152   0.04429957  -7.950 0.00000000000000190 ***
```

```
## yearsq        0.00008847   0.00001112   7.955 0.00000000000000182 ***
## female        0.00362584   0.00291375   1.244               0.213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3303 on 52149 degrees of freedom
## Multiple R-squared:  0.001295,   Adjusted R-squared:  0.001237
## F-statistic: 22.53 on 3 and 52149 DF,  p-value: 0.00000000000001425
```

Old control model

```
## Call:
## lm(formula = very_happy ~ year + yearsq + old, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3697 -0.3222 -0.2938  0.6563  0.7267
##
## Coefficients:
##                   Estimate   Std. Error t value            Pr(>|t|)
## (Intercept) 30.536719485 62.031124120   0.492               0.623
## year        -0.029070263  0.062293915  -0.467               0.641
## yearsq       0.000006973  0.000015639   0.446               0.656
## old          0.044714049  0.004179446  10.699 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4641 on 52149 degrees of freedom
## Multiple R-squared:  0.003185,   Adjusted R-squared:  0.003127
## F-statistic: 55.53 on 3 and 52149 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq + old, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5976 -0.5412  0.4029  0.4452  0.5192
##
## Coefficients:
##                    Estimate    Std. Error t value             Pr(>|t|)
## (Intercept) -370.86907103   66.21333790  -5.601         0.0000000214 ***
## year           0.37193853    0.06649385   5.594         0.0000000224 ***
## yearsq        -0.00009310    0.00001669  -5.577         0.0000000246 ***
## old           -0.05644208    0.00446123 -12.652 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4954 on 52149 degrees of freedom
## Multiple R-squared:  0.004365,   Adjusted R-squared:  0.004308
## F-statistic: 76.21 on 3 and 52149 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq + old, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1604 -0.1317 -0.1208 -0.1096  0.8915
```

```
##
## Coefficients:
##                 Estimate   Std. Error t value           Pr(>|t|)
## (Intercept) 341.33235154  44.14381878   7.732 0.0000000000000108 ***
## year         -0.34286827   0.04433083  -7.734 0.0000000000000106 ***
## yearsq        0.00008613   0.00001113   7.739 0.0000000000000102 ***
## old           0.01172804   0.00297426   3.943 0.0000805140396077 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3303 on 52149 degrees of freedom
## Multiple R-squared:  0.001563,   Adjusted R-squared:  0.001505
## F-statistic:  27.2 on 3 and 52149 DF,  p-value: < 0.00000000000000022
```

Adrian Varallyay
October 5, 2021

Appendix F

Interacted models: yearxfemale

```
## Call:
## lm(formula = very_happy ~ year + yearsq + female + year:female +
##     yearsq:female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3669 -0.3172 -0.3045  0.6744  0.7186
##
## Coefficients:
##                    Estimate    Std. Error t value Pr(>|t|)
## (Intercept)   -144.32347520   92.48352704  -1.561  0.11864
## year             0.14607953    0.09288009   1.573  0.11578
## yearsq          -0.00003688    0.00002332  -1.582  0.11373
## female         385.61670958  124.71100712   3.092  0.00199 **
## year:female     -0.38646715    0.12524024  -3.086  0.00203 **
## yearsq:female    0.00009683    0.00003144   3.080  0.00207 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4645 on 52147 degrees of freedom
## Multiple R-squared:  0.001445,   Adjusted R-squared:  0.001349
## F-statistic: 15.09 on 5 and 52147 DF,  p-value: 0.000000000000007621
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq + female + year:female +
##     yearsq:female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5821 -0.5610  0.4191  0.4367  0.5011
##
## Coefficients:
##                    Estimate    Std. Error t value Pr(>|t|)
## (Intercept)   -342.06229264   98.76922525  -3.463 0.000534 ***
## year             0.34330217    0.09919274   3.461 0.000539 ***
## yearsq          -0.00008599    0.00002490  -3.453 0.000555 ***
## female        -135.13627525  133.18706528  -1.015 0.310284
## year:female      0.13524069    0.13375226   1.011 0.311960
## yearsq:female   -0.00003384    0.00003358  -1.008 0.313588
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4961 on 52147 degrees of freedom
## Multiple R-squared:  0.00161,    Adjusted R-squared:  0.001514
## F-statistic: 16.82 on 5 and 52147 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq + female + year:female +
##     yearsq:female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -0.1580 -0.1315 -0.1196 -0.1110  0.8940
##
## Coefficients:
##                    Estimate    Std. Error t value          Pr(>|t|)
## (Intercept)    487.38576784   65.76037023   7.412 0.000000000000127 ***
## year            -0.48938171    0.06604235  -7.410 0.000000000000128 ***
## yearsq           0.00012287    0.00001658   7.411 0.000000000000128 ***
## female        -250.48043434   88.67570541  -2.825           0.00473 **
## year:female      0.25122646    0.08905201   2.821           0.00479 **
## yearsq:female   -0.00006299    0.00002236  -2.817           0.00484 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3303 on 52147 degrees of freedom
## Multiple R-squared:  0.001479,   Adjusted R-squared:  0.001383
## F-statistic: 15.45 on 5 and 52147 DF,  p-value: 0.000000000000003265
```

Interacted models: yearxold

```
## Call:
## lm(formula = very_happy ~ year + yearsq + old + year:old + yearsq:old,
##     data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3640 -0.3189 -0.2901  0.6486  0.7138
##
## Coefficients:
##                  Estimate    Std. Error t value Pr(>|t|)
## (Intercept)  172.91470211   80.11346669   2.158  0.03090 *
## year          -0.17231962    0.08046096  -2.142  0.03223 *
## yearsq         0.00004300    0.00002020   2.129  0.03329 *
## old         -344.58695507  126.75778043  -2.718  0.00656 **
## year:old       0.34669742    0.12728903   2.724  0.00646 **
## yearsq:old    -0.00008719    0.00003195  -2.729  0.00636 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4641 on 52147 degrees of freedom
## Multiple R-squared:  0.003385,   Adjusted R-squared:  0.003289
## F-statistic: 35.42 on 5 and 52147 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq + old + year:old +
##     yearsq:old, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5993 -0.5423  0.4018  0.4443  0.5166
##
## Coefficients:
##                  Estimate    Std. Error t value     Pr(>|t|)
## (Intercept) -485.56807244   85.51748281  -5.678 0.0000000137 ***
## year           0.48740704    0.08588842   5.675 0.0000000140 ***
## yearsq        -0.00012216    0.00002156  -5.665 0.0000000148 ***
```

```
## old          274.44325992  135.30816674   2.028       0.0425 *
## year:old      -0.27631607    0.13587525  -2.034       0.0420 *
## yearsq:old     0.00006953    0.00003411   2.038       0.0415 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4954 on 52147 degrees of freedom
## Multiple R-squared:  0.004503,   Adjusted R-squared:  0.004408
## F-statistic: 47.18 on 5 and 52147 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq + old + year:old +
##     yearsq:old, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1626 -0.1302 -0.1198 -0.1106  0.8906
##
## Coefficients:
##                 Estimate   Std. Error t value     Pr(>|t|)
## (Intercept) 313.65337032  57.01731335   5.501 0.0000000379 ***
## year         -0.31508742   0.05726463  -5.502 0.0000000377 ***
## yearsq        0.00007916   0.00001438   5.506 0.0000000370 ***
## old          70.14369514  90.21439697   0.778        0.437
## year:old     -0.07038136   0.09059249  -0.777        0.437
## yearsq:old    0.00001766   0.00002274   0.776        0.438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3303 on 52147 degrees of freedom
## Multiple R-squared:  0.001575,   Adjusted R-squared:  0.001479
## F-statistic: 16.45 on 5 and 52147 DF,  p-value: 0.0000000000000002927
```

Interacted models: femalexold

```
## Call:
## lm(formula = very_happy ~ year + yearsq + old:female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3619 -0.3214 -0.3027  0.6672  0.7120
##
## Coefficients:
##                Estimate  Std. Error t value   Pr(>|t|)
## (Intercept) 58.50081616 62.02105231   0.943      0.346
## year        -0.05717482  0.06228368  -0.918      0.359
## yearsq       0.00001404  0.00001564   0.898      0.369
## old:female   0.02374465  0.00489459   4.851 0.00000123 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4645 on 52149 degrees of freedom
## Multiple R-squared:  0.001447,   Adjusted R-squared:  0.00139
## F-statistic:  25.2 on 3 and 52149 DF,  p-value: 0.0000000000000002786
##
## Call:
## lm(formula = pretty_happy ~ year + yearsq + old:female, data = gss)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5852 -0.5617  0.4155  0.4341  0.5166
##
## Coefficients:
##                  Estimate    Std. Error t value           Pr(>|t|)
## (Intercept) -405.05232984   66.21037538  -6.118 0.0000000009563493 ***
## year           0.40629014    0.06649074   6.110 0.0000000010003508 ***
## yearsq        -0.00010174    0.00001669  -6.095 0.0000000011043645 ***
## old:female    -0.03937212    0.00522521  -7.535 0.0000000000000496 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4959 on 52149 degrees of freedom
## Multiple R-squared:  0.002395,   Adjusted R-squared:  0.002338
## F-statistic: 41.74 on 3 and 52149 DF,  p-value: < 0.00000000000000022
##
## Call:
## lm(formula = not_too_happy ~ year + yearsq + old:female, data = gss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1659 -0.1307 -0.1174 -0.1105  0.8906
##
## Coefficients:
##                 Estimate   Std. Error t value            Pr(>|t|)
## (Intercept) 347.55151368  44.09628980   7.882 0.00000000000000329 ***
## year         -0.34911531   0.04428301  -7.884 0.00000000000000324 ***
## yearsq        0.00008770   0.00001112   7.889 0.00000000000000312 ***
## old:female    0.01562747   0.00348000   4.491 0.00000711584802125 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3303 on 52149 degrees of freedom
## Multiple R-squared:  0.001651,   Adjusted R-squared:  0.001594
## F-statistic: 28.75 on 3 and 52149 DF,  p-value: < 0.00000000000000022
```

Adrian Varallyay
October 5, 2021

Appendix G

ANOVA: Female

```
## Analysis of Variance Table
##
## Model 1: very_happy ~ year + yearsq + female
## Model 2: very_happy ~ year + yearsq + female + year:female + yearsq:female
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)
## 1  52149 11255
## 2  52147 11252  2    3.1114 7.2097 0.0007401 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Analysis of Variance Table
##
## Model 1: pretty_happy ~ year + yearsq + female
## Model 2: pretty_happy ~ year + yearsq + female + year:female + yearsq:female
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1  52149 12834
## 2  52147 12834  2   0.61346 1.2463 0.2876
## Analysis of Variance Table
##
## Model 1: not_too_happy ~ year + yearsq + female
## Model 2: not_too_happy ~ year + yearsq + female + year:female + yearsq:female
##   Res.Df  RSS Df Sum of Sq      F   Pr(>F)
## 1  52149 5690
## 2  52147 5689  2    1.0501 4.8127 0.008129 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA: Old

```
## Analysis of Variance Table
##
## Model 1: very_happy ~ year + yearsq + old
## Model 2: very_happy ~ year + yearsq + old + year:old + yearsq:old
##   Res.Df   RSS Df Sum of Sq      F   Pr(>F)
## 1  52149 11232
## 2  52147 11230  2    2.2561 5.2382 0.005313 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Analysis of Variance Table
##
## Model 1: pretty_happy ~ year + yearsq + old
## Model 2: pretty_happy ~ year + yearsq + old + year:old + yearsq:old
##   Res.Df   RSS Df Sum of Sq      F  Pr(>F)
## 1  52149 12798
## 2  52147 12796  2    1.7719 3.6105 0.02705 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Analysis of Variance Table
##
## Model 1: not_too_happy ~ year + yearsq + old
## Model 2: not_too_happy ~ year + yearsq + old + year:old + yearsq:old
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
```

Adrian Varallyay
October 5, 2021

```
## 1  52149 5688.5
## 2  52147 5688.4  2  0.069279 0.3175 0.7279
```

ANOVA: oldxfemale

```
## Analysis of Variance Table
##
## Model 1: very_happy ~ year + yearsq
## Model 2: very_happy ~ year + yearsq + old:female
##   Res.Df   RSS Df Sum of Sq      F      Pr(>F)
## 1  52150 11257
## 2  52149 11252  1    5.0779 23.534 0.000001231 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Analysis of Variance Table
##
## Model 1: pretty_happy ~ year + yearsq
## Model 2: pretty_happy ~ year + yearsq + old:female
##   Res.Df   RSS Df Sum of Sq      F              Pr(>F)
## 1  52150 12838
## 2  52149 12824  1    13.961 56.777 0.00000000000004961 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Analysis of Variance Table
##
## Model 1: not_too_happy ~ year + yearsq
## Model 2: not_too_happy ~ year + yearsq + old:female
##   Res.Df    RSS Df Sum of Sq      F      Pr(>F)
## 1  52150 5690.2
## 2  52149 5688.0  1    2.1995 20.166 0.000007116 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```