**Progress Report, Labour Economics**

**Xavier Friesen and Cal Hofman**

**A replication of Huysse-Gaytandjieva, A., Groot, W., Pavlova, M., & Joling, C. (2015). Low self-esteem predicts future unemployment. *Journal of applied economics*, *18*(2), 325-346.**

**Operationalising:**

* Self-esteem is measured by the Rosenberg Trait Self-Esteem scale. This is a 10-survey question, conducted in the Personality Survey, with 7 answer options ranging from totally disagree to totally agree. Five questions are negatively phrased. In total this will give a score of 10 (low self-esteem) to 70 (high self-esteem).
  + This measurement differs from the original paper: “Have you recently been thinking of yourself as a worthless person?”, for which the scale was 1 (not at all) to 4 (much more). Answers 3 and 4 were coded as low self-esteem as a dummy. However, the Rosenberg Self-Esteem scale is generally preferred as a measurement of self-esteem in the data, and thus it appears logical to use this measurement since it is available. Other operationalisations might be used as robustness checks.
* We also create a dummy for becoming unemployed. This equals 1 if a person is employed at tx, and lost his job at tx+1. Losing their job is defined as all those that already have had a job, and are actively looking for one (belbezig = 4/6). Thus, students, retirees and people working in the household are not considered unemployed, but are instead excluded as NA.
* Other control variables were obtained from the background survey (e.g. partner status, age, gender, education and number of children), schooling and working survey (job satisfaction, sector and type of work) and health survey (drinking and mental health). The data was obtained yearly, and for the background survey we used the month of July.

**Data Preparation and Descriptive Statistics:**

* Adding extra controls from different surveys increases the number of missing observations significantly. Clearly, not all people respond to the different surveys every year. This is something to keep in mind.
  + A more specific problem is that not everybody responds to the personality survey every year. Thus, we do not have access to the critical independent variable for every person. Instead, we opted to input a Rosenberg score if missing by using the last available measurement for that person. We argue this compromise is acceptable, since the measured trait self-esteem should be relatively constant over time. This assumption is further strengthened by the low within-person variance of the Rosenberg-Score over time.

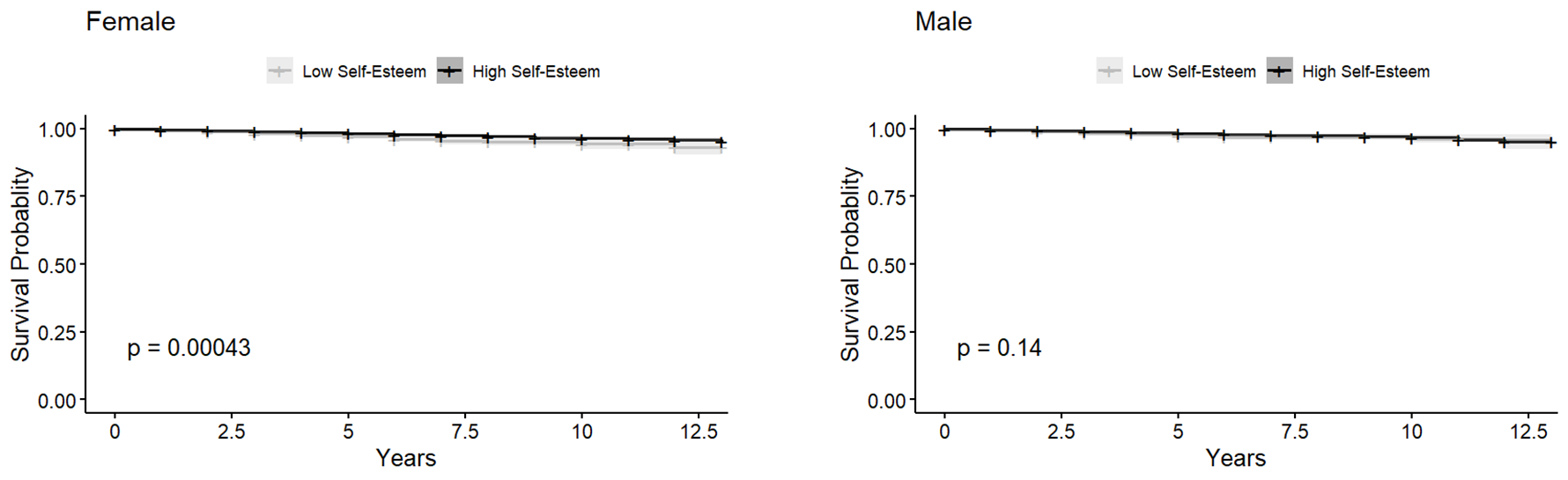


Figure 1: Kaplan Meier Curves Using Rosenberg = 40 as Cut-Off

* The Kaplan Meier curves in Figure 1 show similar results to the original paper. Females with low self-esteem are significantly more likely to lose their job over time. This effect is not statistically distinguishable for males. However, the magnitude of the effect is much smaller than reported graphically in the original paper.
* Additionally, Table 1 and Figure 2 show that very few people with very low self-esteem scores (below 30) experienced a lay-off. While in the original research in the UK about 1 in 7 participants experienced a lay-off, this percentage is much lower in our dataset (7% chance of being laid-off). This signals potentially problems with the power of the data to find (non-linear) relationships between a very low self-esteem and unemployment.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Female** | | **Male** |
| Low Self-Esteem | 11.73% | | 7.98% |
| High Self-Esteem | 88.26% | | 92.02% |
|  | **Not Laid Off (Number of Respondents)** | **Laid Off (Number of Respondents)** | **Total Number of Respondents** |
| Low Self-Esteem | 937 | 85 | 1022 |
| High Self-Esteem | 8381 | 567 | 8948 |
| Total | 9318 | 652 | 9970 |

Table 1: Self-Esteem Rates by Gender and Lay-Off Using Rosenberg-40 as Cut-Offs

**A comparison of graph and graph

Description automatically generated with medium confidence**

Figure 2: Number of Lay-Offs per Rosenberg Score

**Empirical Strategy:**

* We use a very similar approach to the original authors. We use a piecewise constant regression model using covariates at year t, to predict unemployment status at year t+1. By using this concept of Granger Causality, we control for any possible endogenous effects of self-esteem decreasing due to becoming unemployed. Additionally, using survival analysis makes it possible to include participants which dropped out, or have otherwise missing observations in some of the years.
* As characteristic of a piecewise model, we assume the odds of becoming unemployed are constant within the year, but are allowed to vary over time.
* We control for general time effects in all models, to allow for the introduction of new participants in the survey at different times, and control for general economic shocks.
* In addition, we included several covariates in the models, and uses the Akaike Information Criterion to identify which model is preferred. The covariates match closely the covariates identified by the original authors, except that the categories might be slightly different. We could not find data on regional unemployment effects or union participation. We would argue both are less important in the context of the Netherlands compared to the UK.
* We did include an additional control of job satisfaction in Model G not included in the paper, which has clear theoretical impact on becoming unemployed. We also attempted to include a control for how much a person’s health was hindering their specific job. However, including this covariate lead to a violation of the proportional hazards of the model, and was thus impossible. As identified as a problem by the original authors, we also do not have data on job performance.

**Identifying the Model**

In Model A in Table 3, a base model without covariates, we find a strong statistically significant effect of self-esteem on the probability of getting laid-off. Using the mid-score for Rosenberg of 40 as a cut-off for the dummy of having low self-esteem, we find that females with low self-esteem are 44% more likely to lose their job. This is a slightly larger effect as found in the original study, and more importantly, this effect is also statistically significant for males in contrary to the original results.

Adding controls for age, education, children and partner status do not alter these conclusions, although these covariates do improve the fit of the model. After adding dummies for sector of employment and type of work, the effect of self-esteem is no longer statistically significant for males, and the results are thus very similar compared to the original paper.

However, the results for Model F differ sharply from the original paper. When including dummies for heavy alcohol consumption and mental health, the effects of self-esteem on the chance of unemployment rates disappears. This in sharp contrast to the original paper, for which the results were robust to the inclusions of the same covariates. Seemingly, it’s not self-esteem which matters in the data, but heavy alcohol consumption and mental health status which drive the causal effect. Finally, in Model G we add job satisfaction, not included by the original paper, which was statistically significant and improved the model.

The original authors used Wald Hypothesis Testing to choose between which model was best. However, since the number of observations differs sharply between models due to attrition rates in the LISS-surveys, this is not possible. Instead, we use the AIC-criterion, which strongly prefers Model G. Using the Bayesian Information Criterian would not have altered the models preference. Model G is also preferred if analysing other cut-off values for the dummy or when using the continuous Rosenberg Scores (tables not reported), and will thus be used in the further analysis.

**Estimation Results**

In Table 2, the preferred speciation of model G is used to estimate the effects of self-esteem for different cut-off values. Only when choosing 35 as a cut-off value the effect of self-esteem on job unemployment is statistically significant for females. For males, the effect is never statistically significant. Due to lack of power in the data, standard errors for Model A are huge, and suggest Model A would never be able to show a statistically significant result.

We will need to further research whether the specific effect at 35 is an anomaly or perhaps due to the fallacy of multiple hypothesis testing or whether this signals a non-linear relationship.

Model E in Table 2 shows there is no linear effect of Rosenberg Score on the chance of becoming unemployed controlling for all other covariates.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **(A)** | **(B)** | **(C)** | **(D)** | **(E)** |
| Cut-off Value | **Rosenberg-25** | **Rosenberg-35** | **Rosenberg-40** | **Rosenberg-45** | **Linear Rosenberg-Score** |
| **A. Women** |  |  |  |  |  |
| Hazard Ratio Self-Esteem | 0.677 | 0.585\* | 0.932 | 1.091 | 1.0034 |
| Standard Error of Self-Esteem | 0.693 | 0.253 | 0.203 | 0.606 | 0.0067 |
| *Goodness of Fit* |  |  |  |  |  |
| Log-likelihood | -1575 | -1573 | -1652 | -1575 | -1575 |
| Deviance | 154 | 157 | 153.3 | 153 | 153 |
|  |  |  |  |  |  |
| **B. Male** |  |  |  |  |  |
| Parameter of Self-Esteem | 0.271 | 1.377 | 1.309 | 0.778 | 0.9913 |
| Standard Error of Self-Esteem | 0.685 | 0.467 | 0.289 | 0.188 | 0.0076 |
| *Goodness of Fit* |  |  |  |  |  |
| Log-likelihood | -1484 | -1485 | -1596 | -1484 | -1485 |
| Deviance | 223 | 221 | 222.1 | 223 | 223 |
| **Notes:** a) \*\*\* = .000, \*\* = .01, \* = .05. b) All regressions are estimated using the specification of model G. Standard errors are robust. No covariates were significant in a Schoenfeld Residuals test. | | | | | |

Table 2: Estimation Results for Different Cut-Offs

**Next empirical steps:**

* Further researching a possible non-linear effect
* Checking if the participants that failed to fill in the health survey is different from the other participants, in which case attrition might be the reason there is no effect.
* Further robustness checks by using different measures for self-esteem, and perhaps slightly different model specifications
* Now, unemployment is only being recorded if a person is still unemployed at the next year. If somebody loses their job, and finds a new job before the next wave of the survey this is not visible in the data. This was also a not mentioned problem with the original research. We can try to merge to calculate unemployment data by month, and re-do the analysis to increase the sample size and precision of the analysis

**Table 3: Fitting Covariates for Piecewise Constant Model (using Rosenberg = 40 as cut-off)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **(A)** | **(B)** | **(C)** | **(D)** | **(E)** | **(F)** | **(G)** |
| **A. Women** |  |  |  |  |  |  |  |
| Hazard Ratio of Self-Esteem | 0.561\*\*\* | 0.541\*\*\* | 0.560\*\*\*1 | 0.638\*\*\* | 0.659\* | 0.827 | 0.932 |
| Robust Standard Error of Self-Esteem | 0.139 | 0.141 | 0.141 | 0.167 | 0.173 | 0.189 | 0.203 |
| Number of Participants | 24318 | 24318 | 24318 | 20138 | 19298 | 16889 | 15238 |
| Number of Lay Offs | 369 | 369 | 369 | 307 | 302 | 273 | 248 |
| *Goodness of Fit* |  |  |  |  |  |  |  |
| Log-likelihood | -2598 | -2598 | -2598 | -2112 | -2068 | -1836 | -1652 |
| Deviance | 15.1 | 20.5 | 30.2 | 57.6 | 80.7 | 96.7 | 153.3 |
| Number of Parameters | 1 | 3 | 5 | 13 | 26 | 28 | 29 |
| Akaike Information Criterion | 5183 | 5181 | 5175 | 4193 | 4108 | 3631 | 3208 |
|  |  |  |  |  |  |  |  |
| **B. Male** |  |  |  |  |  |  |  |
| Hazard Ratio of Self-Esteem | 0.631\* | 0.604\*\* | 0.653\* | 0.691 | 0.679 | 0.976 | 1.309 |
| Robust Standard Error of Self-Esteem | 0.181 | 0.185 | 0.184 | 0.205 | 0.213 | 0.246 | 0.289 |
| Number of Participants | 22852 | 22852 | 22852 | 19204 | 18514 | 16506 | 15247 |
| Number of Lay Offs | 340 | 340 | 340 | 298 | 292 | 267 | 238 |
| *Goodness of Fit* |  |  |  |  |  |  |  |
| Log-likelihood | -2388 | -2388 | -2388 | -2055 | -2005 | -1806 | -1596 |
| Deviance | 5.7 | 31.1 | 64.5 | 94.2 | 158.8 | 173.2 | 222.1 |
| Number of Parameters | 1 | 3 | 5 | 13 | 26 | 28 | 29 |
| Akaike Information Criterion | 4773 | 4752 | 4722 | 4042 | 3903 | 3495 | 3028 |
| **Notes**: a) \*\*\* = .000, \*\* = .01, \* = .05. b) Model A includes no covariates, in model B age and education are added, model C includes children and partner status, model D includes dummies for sector of employment, model E includes dummies for type of job and model F includes dummies for mental health and excessive alcohol consumption. In Model G, not included in the original paper, a measure of job-satisfaction is included. Unfortunately there was no data available on union participation or regional unemployment rates, as included in the original paper. c) all standard errors are robust, d) using BIC instead of AIC would give the same conclusion. 1 Partner is statistically significant in a Schoenfeld Residuals Test, indicating that for this regression the proportional hazards assumption might be violated. | | | | | | | |
|  |
|  |
|  |