[Re] Marriage and happiness: Providing evidence against a relationship between inequality and happiness in Oishi, Kesebir, and Diener (2011)

Lauren Heintz, Tara Wilson, Ben Mathieu, Will Wright

01 March, 2020

A reference implementation of:

Which is a reference implementation of:

There is a crisis of reproducibility and replication in the data science field. In an attempt to bolster confidence in scientific results, the following paper is a replication of "Marriage and happiness: Providing evidence against a relationship between inequality and happiness in Oishi, Kesebir, and Diener (2011)" (Grunberg, Kim, and Kim 2014a). This paper follows the ReScience replication method(Rougier and contributors 2020) which strongly encourages open source practices as underlined by the format of this project: a github compendium. This particular replication is unique in that the paper being replicated is an alternate implementation of the paper:"Income Inequality and Happiness" (Oishi, Kesebir, and Diener 2011). While the first paper draws a connection between happiness and income, the second calls into question the data used to obtain the results stated, and pulls in additional data to re-assess the analysis. This second analysis, which focused on the relationship between marriage and happiness, is the subject of replication in this paper.

Introduction

In the spirit of reproducible science, a replication of the paper "Marriage and happiness: Providing evidence against a relationship between inequality and happiness in Oishi, Kesebir, and Diener (2011)" (Grunberg, Kim, and Kim 2014a) is conducted according to the ReScience standards. ReScience defines replication as:

"a scientific study (computational or other) means repeating a published protocol, respecting its spirit and intentions but varying the technical details. For computational work, this would mean using

different software, running a simulation from different initial conditions, etc. The idea is to change something that everyone believes shouldn't matter, and see if the scientific conclusions are affected or not" (Rougier and contributors 2020).

This replication is unique because it is a replication of a replication. The first paper, "Income Inequality and Happiness" (Oishi, Kesebir, and Diener 2011), is a study on happiness and specifically focuses on income equality. The study makes the conclusion that in years where there is greater income inequality in America, people are less happy. The paper attributes this negative relationship to meta variables such as perceived trust and perceived fairness in society. The second paper (Grunberg, Kim, and Kim 2014a) calls into question the data and the conclusions used by this study. The second paper replicates the first factoring in additional data like race, gender, and marriage to test for possible confounding variables. Most noteably, a relationship between marital status (single vs married) results in the only significant factor of a linear regression using all variables. In essence, the second paper makes the claim that when controlling for marriage, there is no relationship between income inequality and the happiness of Americans - invalidating the first paper's results.

This paper will now be a second replication, focused on critiquing the methods and traceability of the second paper. Throughout the analysis, some critical pieces of missing data led to challenges in reproducing the results. In an effort to reproduce results, we had to find our own data and sources to tie in for the analysis.

Methods

Overview

The paper of interest(Grunberg, Kim, and Kim 2014b) used linear models to show that the correlations drawn by the original findings(Oishi, Kesebir, and Diener 2011) do not hold once the interactions of variables known to be correlated with the response variables are added to the model.

In this replication paper, three artifacts from (Grunberg, Kim, and Kim 2014b) were reproduced. The first artifact, Figure 3., plots the mean happiness by year and race, the second, Table 1, shows the linear regression model's coefficients and the third, Table 2, shows the correlation coefficients of the different independent variables used in the linear mixed-effect models created.

Dataset

Marriage and Happiness(Grunberg, Kim, and Kim 2014b), obtained data from the General Social Survey (GSS) and the US Census Bureau. The subset of GSS data provided omitted the RACE and MARITAL columns crucial to the replication of the figures and models. In order to complete the missing data, the original dataset was queried directly using the gssr(Healy 2019) R package.

The dataset from gssr contained a larger time interval than the original paper. It was filtered back to match the original interval so that only years prior to 2008 are used in both analyses. The new dataset was subsequently merged with the original US Census Bureau data found in Gini_families.csv(Grunberg, Kim, and Kim 2014b) by appending the "total" column to the GSS data using "year" as the join key.

Several encodings for the data were updated in the original paper which were replicated in order to allow accurate comparisons. The "happy" variable was re-encoded in a reversed scheme so that the numerical factor would reflect an increasing level of happiness. The income data was re-encoded on a log scale. Race was encoded numerically where "White" was set to 1. The "married" variable was also set to numeric encodings with married respondents denoted by those answering 1.

Alterations

Aside from gathering the input data from a new source, several different packages were used to ensure the reproduced results were adequately tested. These will be discussed further in their respective sections below.
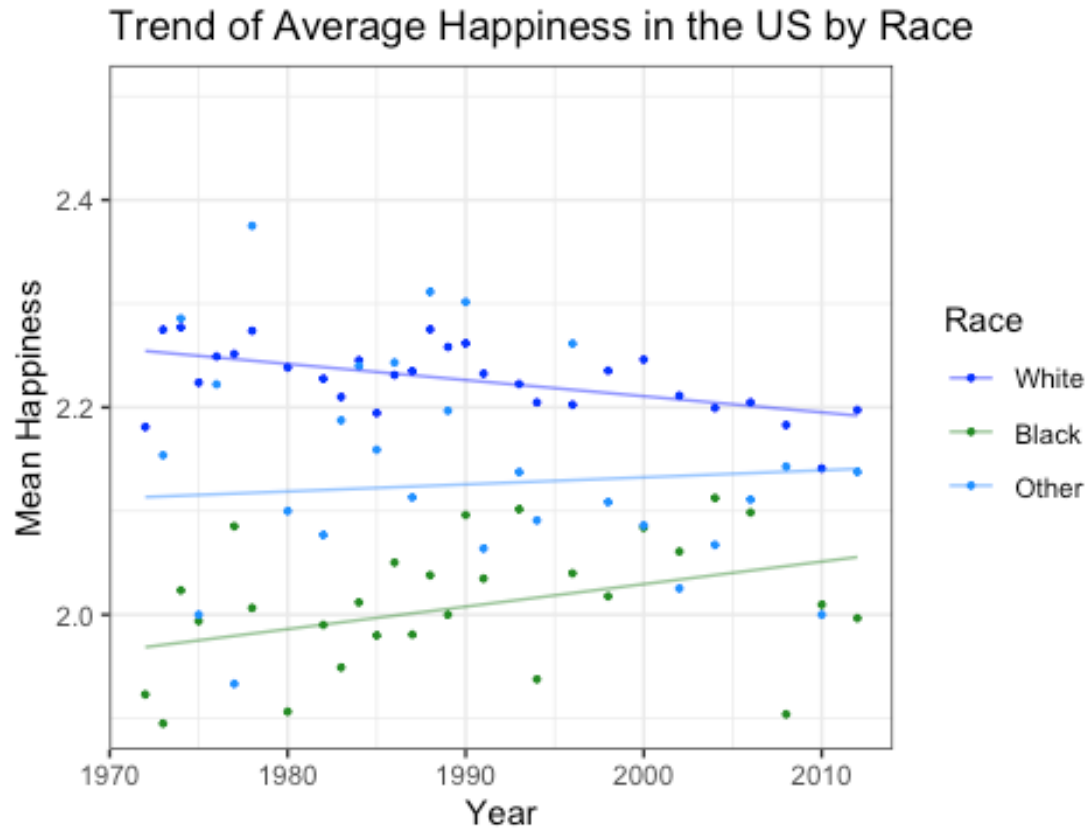
Results

Figure 3

Replicating Figure 3

Figure 3 in the original publication plots mean happiness of respondents over time. This data is categorized into 3 groups by racial identity: White, Black and Other. Best-fitting regression lines are included to display the trend of mean happiness for the 3 populations.

Figure 1: Figure 3 from the original paper
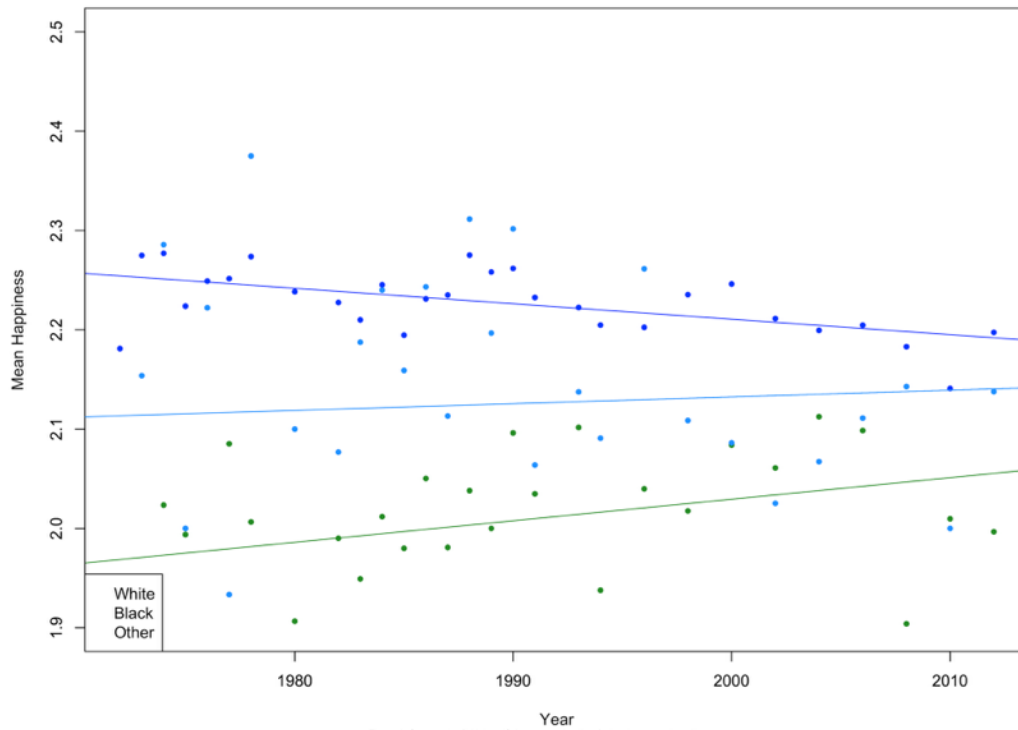
The original R code used the standard R libraries for visualizations. The ggplot2 package was leveraged to replicate the figure with the new dataset. Axis labels, lines and colors were adjusted to match the original output.

Replicating Table 1

Table 1 in the original publication explored the correlations between happiness and various control variables found in the dataset.

	Estimate	p value
Logged Income	0.19	0.00
Age	0.03	0.00
Sex	0.01	0.23
White	0.12	0.00
Married	0.23	0.00

Table 1: Correlations between happiness and control variables

Figure 2: Table 1 from the original paper

The R function cor.test() was used to get the Pearson correlation coefficient between the different independent variables (income on a log scale, age, sex, race, married) and the response variable, happiness. Grundberg et al.'s code was slightly modified to keep the same encoding of variables given the new dataset input.

Figure 3: Replication of Table 1

Figure 3: Replication of Table 1

We were able to almost exactly reproduce the results from Table 1 in the original paper as shown in Figure 2. The only difference being that the estimate for Sex was 0.01 in the publication compared to 0.00 in our replication. This may be due to rounding versions or an underlying exclusion of outliers we saw with the above plot.

The steps taken to reproduce this table are identical to those in the original publication aside from a new input dataset and the use of a combination of the kableExtra and xtable R packages to produce the output compared with xtable alone used in the original publication.

Replicating Table 2

Table 2 in the original publication explores the output of various linear models on the dependent variable, happiness. This is central to the message of the paper and was therefore critical to investigate.

Table 2: Multilevel Linear Models

	Dependent variable:					
	Happiness					
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	−0.385**	−0.461**	−0.413**	−0.386**	−0.177	0.332
	(0.194)	(0.189)	(0.195)	(0.194)	(0.175)	(0.210)
REALINClog		0.125***				
		(0.003)				
Age			0.001***			
			(0.0002)			
factor(SEX)2				0.007		
				(0.006)		
White					0.188***	
					(0.008)	
Married						0.298***
						(0.006)
Constant	2.351***	1.140***	2.306***	2.348***	2.112***	1.899***
	(0.077)	(0.081)	(0.078)	(0.077)	(0.070)	(0.084)
Observations	48,318	43,564	48,318	48,318	48,318	48,318
Log Likelihood	−46,611.120	−41,048.910	−46,590.320	−46,614.720	−46,308.240	−45,295.860
Akaike Inf. Crit.	93,230.250	82,107.830	93,190.630	93,239.430	92,626.470	90,601.730
Bayesian Inf. Crit.	93,265.390	82,151.240	93,234.560	93,283.360	92,670.400	90,645.650

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 4: Table 2 from the original paper

Grundberg et al used the lme4 package to generate coefficients of linear mixed-effects models using the intercept to capture year as the random effect (see (Bates et al. 2015) section 2).

===
===================== Dependent variable:
——————————————————————————————————— Happiness
(1) (2) (3) (4) (5) (6)
——————————————————————————————————— - Gini -0.385**
-0.461** -0.424** -0.386** -0.177 0.331
(0.194) (0.189) (0.196) (0.194) (0.175) (0.210)

REALINClog 0.125***
(0.003)

Age 0.001***
(0.0002)

factor(sex)2 0.007
(0.006)

White 0.188***
(0.008)

Married 0.298***
(0.006)

Constant 2.351*** 1.140*** 2.309*** 2.348*** 2.112*** 1.900***
(0.077) (0.081) (0.078) (0.077) (0.070) (0.084)

Observations 48,318 43,564 48,158 48,318 48,318 48,306 Log Likelihood
-46,611.120 -41,048.920 -46,419.990 -46,614.720 -46,308.240 -45,282.940
Akaike Inf. Crit. 93,230.250 82,107.830 92,849.970 93,239.430 92,626.470
90,575.880 Bayesian Inf. Crit. 93,265.390 82,151.240 92,893.890 93,283.360
92,670.400 90,619.810

===
================================== Note: *p<0.1; **p<0.05;* p<0.01

x

==

Dependent variable:

Happiness

	(1)	(2)	(3)	(4)	(5)	(6)
Gini	-0.385**	-0.461**	-0.424**	-0.386**	-0.177	0.331
	(0.194)	(0.189)	(0.196)	(0.194)	(0.175)	(0.210)
REALINClog						0.125***
						(0.003)
Age						0.001***
						(0.0002)
factor(sex)2						0.007
						(0.006)
White						0.188***
						(0.008)
Married						0.298***
						(0.006)
Constant	2.351***	1.140***	2.309***	2.348***	2.112***	1.900***
	(0.077)	(0.081)	(0.078)	(0.077)	(0.070)	(0.084)

Observations	48,318	43,564	48,158	48,318	48,318	48,306
Log Likelihood	-46,611.120	-41,048.920	-46,419.990	-46,614.720	-46,308.240	-45,282.940
Akaike Inf. Crit.	93,230.250	82,107.830	92,849.970	93,239.430	92,626.470	90,575.880
Bayesian Inf. Crit.	93,265.390	82,151.240	92,893.890	93,283.360	92,670.400	90,619.810

==

Note: *p<0.1; **p<0.05; ***p<0.01

Again we were able to almost identically reproduce the results from Table 2 in the original paper, shown in Figure 4. Aside from slight variability in the second and third decimal places, our results from the linear models are very close. There are several differences with the number of observations, listed in the bottom of the table, which is to be expected as we are using a package with the GSS data and it is likely not an identical snapshot to the one used in the original paper. However, this variability is minor in scale. Even though slight variabilities were observed in different cells throughout this table, no disparities were large enough to impact the significance of the results as denoted in the tables with asterisks. Therefore, even with new data sources we found that the conclusions hold for the linear model analysis.

Again, the steps taken to reproduce this table were nearly identical as the original paper aside from a new dataset source and the use of the kable and stargazer packages to produce the output as opposed to strictly stargazer LaTeX table results.

Conclusion

Replication of this paper was ultimately contingent on tracing back the original dataset. Without this resource being cited and publicly available, it would not have been possible to replicate the different tables and figure presented in the original paper. The lack of necessary information included in the published data provided for replication is something that should likely be addressed by the authors as this was a major deterrent to the goal of reproducibility.

Under conditions of both a new source dataset and different R packages used for data manipulation and visualization, the 3 exhibits in the original publication were able to be replicated. With only minor discrepancies in the numbers themselves, the conclusions drawn from the paper's graphics were not significantly altered. Successful replication of this paper is largely a result of transparency in the original publication providing clarity around data encoding and the models used for analysis. This provided a platform upon which we were able to adapt the code and generate results that in large part align directly with those in the publication.

References Cited

Reference 1: Grunberg, Kim, and Kim (2014a) Reference 2: Grunberg, Kim, and Kim (2014b) Reference 3: Oishi, Kesebir, and Diener (2011) Reference 4: Healy (2019) Reference 5: Rougier and contributors (2020) Reference 6: Bates et al. (2015)

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software, Articles*. https://doi.org/10.18637/jss.v067.i01.

Grunberg, Rebecca L., Hyejun Kim, and Minjae Kim. 2014a. "Marriage and Happiness Grunberg Kim Kim.pdf." Harvard Dataverse. https://doi.org/10.7910/DVN/25655/MB980L.

———. 2014b. "Replication data for: Marriage and happiness: Providing evidence against a relationship between inequality and happiness in Oishi, Kesebir, and Diener (2011)." Harvard Dataverse. https://doi.org/10.7910/DVN/25655.

Healy, Kieran. 2019. "gssr: General Social Survey data for use in R." http://kjhealy.github.io/gssr.

Oishi, Shigehiro, Selin Kesebir, and Ed Diener. 2011. "Income Inequality and Happiness." *Psychological Science* 22 (9): 1095–1100. https://doi.org/10.1177/0956797611417262.

Rougier, Nicolas P., and contributors. 2020. "ReScience C: Overview of the submission process." http://rescience.github.io/write/.