Can Income Raise Individual's Happiness?*
A Analysis of Income Rank and Well-Being

Qingya Li

27 February 2022

Abstract

The report replicates the paper "The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment" by Perez-Truglia (2020). The author estimates the correlation between income transparency and Norwegian individual subjective-well being. He builds difference in difference models by importing the German dataset and he finds the actual disclosure effect in Norway. He concludes the subjective well-being gap between the poor and the rich increases as the income transparency increases. The replicative results are similar to the original. We find individual income rank associates about 0.3 and 0.6 increase with their happiness and life satisfaction, respectively.

Contents

---

1 Introduction

Whether income can raise an individual's happiness has been a controversial topic for many years. Economists and policymakers are interested in studying this topic due to its central importance. Many researchers believe in a significant relationship between individual income and individual happiness. For instance, the paper published by Clark, Frijters, and Shields(2008) studied the relative income term in the utility function, happiness, and utility and how does relative income affect other economic behaviors by conducting the models. They found that individual compares income to others or oneself (E.Clark, Frijters, and Shields 2008). Income is also positively correlated to individual happiness (E.Clark, Frijters, and Shields 2008). They concluded that we need to take relative income seriously.

However, other researchers disagree that income does not play a significant role in well-being. According to the Esterlin Paradox, happiness does not rise as income continues to increase over time. By studying several countries' time series data, he summarized some conclusions: The United States case against the idea that national income can increase happiness (Esterlin 2005). Also, even though Japan is a developed country with a high national income rank, its happiness rank is the third from the bottom among 20 countries (Esterlin 2005).

This report replicates the paper written by Perez-Truglia (2020). As tax records became accessible online in 2001, in Norway, people are permitted to review their income as well as any other people (Perez-Truglia 2020). The author analyzes how does income transparency affects Norwegians' well-being. Perez-Truglia (2020) first create histograms and line charts to compare the google search volumes of each keyword in Norway. He finds Norwegians are interested in searching tax lists, typically when the tax agency realizes the income data. Then he conducts difference in difference models. To limit the time effect, he uses German data as a counterfactual, which does not have disclosure in 2001 (Perez-Truglia 2020). By comparing the subjective well-being between Norway and Germany, Perez-Truglia (2020) can find the real discolored effect on subjective well-being in Norway. The respondent variables in the models are happiness and life satisfaction , respectively. He purposes that higher-income transparency increases 29% of the happiness gap between poor and rich individuals, and 21% life satisfaction gap. His conclusion is relative to the study by Clark, Frijters, and Shields (2008). People compare their income to others and which may lead to the rich well-being becoming higher and higher, the poor situation is converse.

We replicate the original plots and models by using the analytic dataset from the paper. We estimate that the respondents' income rank approximately raise 0.3 and 0.6 their happiness and life satisfaction, respectively (p values are less than 0.001). Meaning higher individual income rank is associated with higher individual happiness. Additionally, other factors such as gender, education, age can also affect subjective well-being.

2 Data

The report is written by the **R** statistical language (R Core Team 2020), the packages of **tidyverse** (Wickham et al. 2019) and **dplyr** (Wickham et al. 2021). Stata files are read by **haven** packages (Wickham and Miller 2020). In the report, we use **RMarkdom** (Allaire et al. 2020) and **bookdown** (Xie 2021) to output the results. The graphs are created by **ggplot2** (Wickham 2016).

The original paper[1] uses multiple sources which are in .dta format (Perez-Truglia 2020). The codes are run by Stata. The study is mainly based on the dataset of Norwegian Monitor Survey, google searches by country, google searches over time, name browsing skattelister, and GSOEP. The GSOEP dataset is available at the German Socio-Economic Panel[2] which you need to require permission to access.

The Norwegian Monitor data is collected through a cross-sectional survey (the survey is self-response by mail) by the research institute Ipsos MMI (Perez-Truglia 2020). Thus, the population is Norwegians and the sample is the one who completes the survey. However, non-response bias may exist. The survey questions

---

[1]The original paper is available at: https://doi.org/10.1257/aer.20160256

[2]The GSOEP dataset is available at: https://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/GSOEP

focus on Norwegians' subjective well-being, which is measured by happiness and life satisfaction (Perez-Truglia 2020). The dataset contains the information from every other year in 1985-2013 (Perez-Truglia 2020). The respondence date is not recorded but the researchers consider respondents finished the survey during September and December (Perez-Truglia 2020). As Perez-Truglia (2020) mentioned, the tax data for the previous year is released in October by the agency. Therefore, the respondents complete the survey when the income transparency is most significant, which may overestimate the effect of subjective well-being. Meanwhile, the survey in 2001 may finish before the tax record is accessible online which may lead to an underestimation (Perez-Truglia 2020). There are 49,009 observations in the dataset are used in my report.

The google searches by country and the google searches over time data are collected by Google Trends, the datasets show the popularity of income searches in 2010(Perez-Truglia 2020). Overall, four keywords are skatterlister, youtube, weather, and tax. Specifically, skatterlister means "tax list", it refers to skattlister and skattelistene (Perez-Truglia 2020). We will normalize the search keywords to a fraction of Youtube (Perez-Truglia 2020) and compare each search volumes in Figure 1, which contains a robustness check by using Sweden data. We also uses line chart to show the number of Google search for each categories of keywords in Norway between 2010 and 2011 in 2.

Figure 1 illustrates each of the annual searches volumes, relative to Youtube, between Norway and Sewden
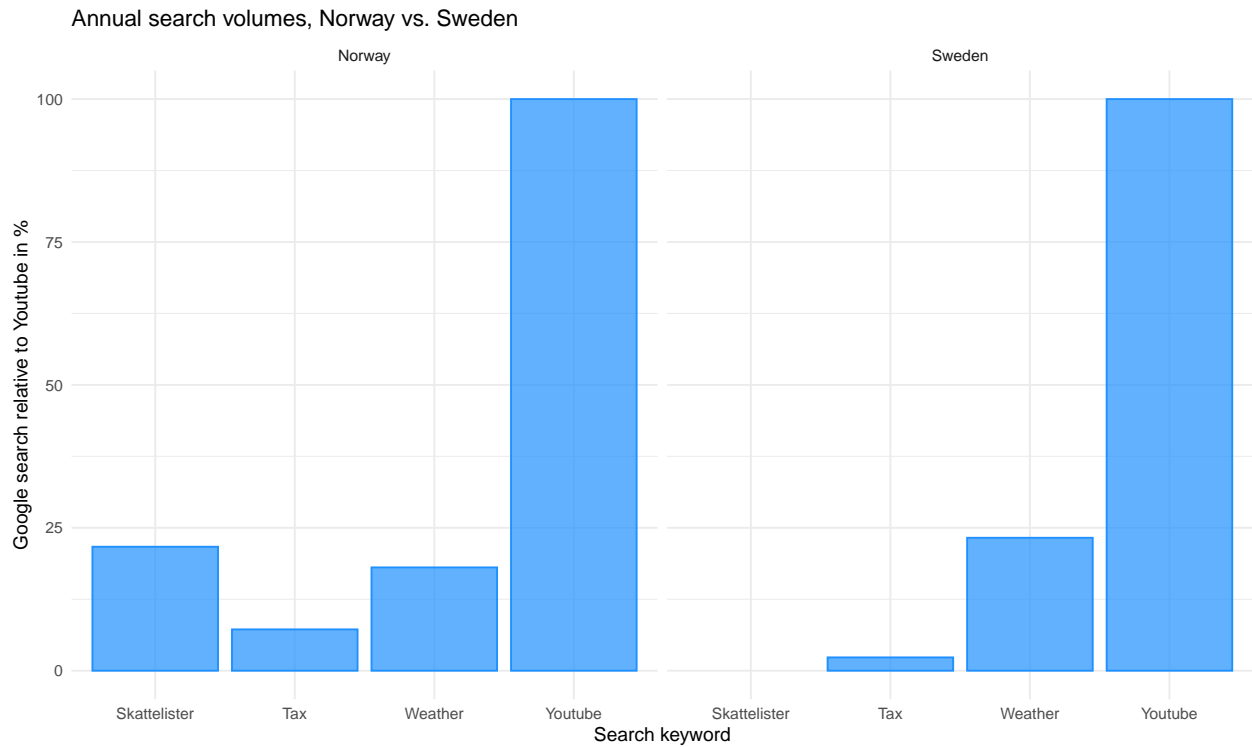

Figure 1: The annual numbre of Google search for different categories of keywords, relative to the search for Youtube

The bar plot shows that Norwegians are more likely to search weather than tax, and the times they search tax lists are much more than taxes. The evidence proves the uses of tax lists may not correlate to the taxes. Compare to Swedes, Norwegians are more interested in searching taxes but less interested in the weather. Overall, the search volumes for weather and Youtube are most likely between the two countries.

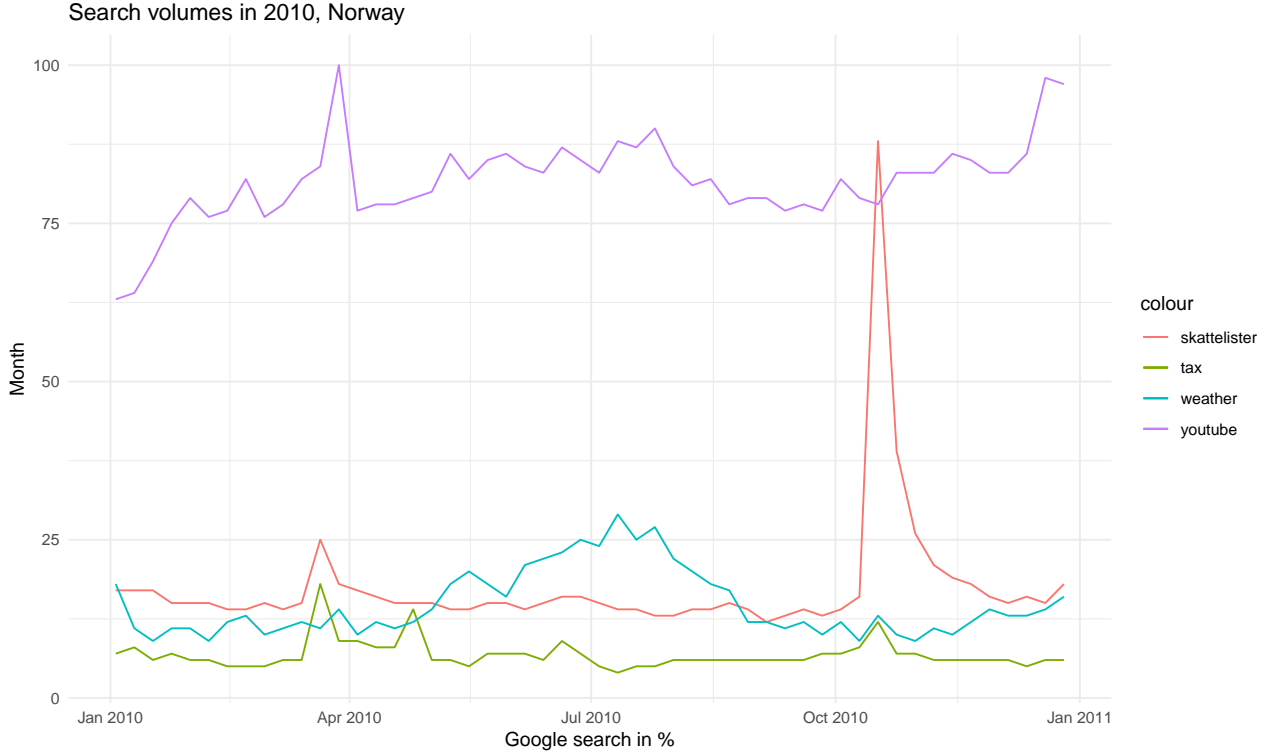Figure 2 demonstrates the trend of the number of Google search for the keywords in Norway in 2010-2011.

Figure 2: The number of Google search for each categories of keywords in Norway from Jan, 2010 - Jan, 2011

The google search for the tax lists is almost twice the search for taxes, and the search for weather is similar to the search for taxes. We notice that the search for tax lists suddenly raise a lot during October and even more than the search for Youtube. As we mentioned above, the tax agency released the tax data of the previous year every October which may cause an increase in searching tax lists. The evidence proves Norwegians are more interested in tax lists than other events.

3 Model

According to the original paper (Perez-Truglia 2020), we intend to find the factors that may affect the subjective well-being of people by conducting difference in difference models. We predict that higher income transparency is associated with higher subjective well-being.

The replication baseline models are presented as below:

$$(1) SWB_{i,t} = \alpha_1 IncomeRank_{i,t} + \alpha_2 IncomeRank_{i,t} I_t^{01-13} + X_{i,t}\beta + \delta_t + \epsilon_{i,t}$$

- $SWB_{i,t}$ is subjective well-being measurment, where the higher scores of SWB meaning higher well-being that individual feels
- $IncomeRank_{i,t}$ measures the individual's position in the national distribution of household income in a year, from 0 to 1. 0 represents individual gain the lowest income in the distribution
- I_t^{01-13} is a dummy variable of income transparency. 1 is high income transparency, 0 is low, at the period of 2001-2013
- $X_{i,t}$ are the controls variables

- δ_t is effects by years
- $\epsilon_{i,t}$ is random error term
- i represents each individual and t represents the year

$$(2) SWB_{i,t} = \alpha_1 IncomeRank_{i,t} + \alpha_2 IncomeRank_{i,t}I_t^{01-13} + \gamma IncomeRank_{i,t}(t-1985) + X_{i,t}\beta + \delta_t + \epsilon_{i,t}$$

- γ is the linear trend of the income and happiness between 1985 and 2013.

$$(3) SWB_{i,t} = \alpha_1 IncomeRank_{i,t} + \alpha_2 IncomeRank_{i,t}I_t^{01-13} + \alpha_3 IncomeRank_{i,t}I_t^{97-00} + X_{i,t}\beta + \delta_t + \epsilon_{i,t}$$

- I^{97-00} is a dummy variable, it is equal to 1 if at the period of 1997-2000, 0 otherwise. It indicates the income rank just before the agency released income transparency. Assume the income transparency significantly increase the subjective well-being, we estimate that the α_2 is positive and α_3 is 0 meaning the period 1997-2000 does not affect the subjective well-being.

$$(4) SWB_{i,t} = \alpha_1 IncomeRank_{i,t} + \alpha_2 IncomeRank_{i,t}I_t^{01-13} + \alpha_3 HigherInternet_{i,t}$$

$$+\alpha_4 HigherInternet_{i,t}I_t^{01-13} + \alpha_5 IncomeRank_{i,t}HigherInternet_{i,t}$$

$$+\alpha_6 IncomeRank_{i,t}HigherInternet_{i,t}I_t^{01-13} + X_{i,t}\beta + \delta_t + \epsilon_{i,t}$$

- $HigherInternet_{i,t}$ is a dummy variable, it is 1 if individual has above-median internet access at home. The original paper use the triple-difference specification to improve the model (Perez-Truglia 2020). The author consider there may exist factors other than income transparency will affect the subjective well-being (Perez-Truglia 2020). Therefore, he creates this variable from internet access data to control the effects.

4 Results

In order to replicate the results, we decide to run Ricardo's codes to generate the analytic dataset in **Stata** (the analytic dataset named "norway_data_ready_to_be_analyzed"), instead of using the raw Norwegian Monitor Survey dataset.

Table 1: Summary statistics in mod1

Variable	Estimate	P_value
(Intercept)	1.917e+00	0.404691
imp_hh_rank	3.173e-01	< 2e-16
post_2001	-7.563e-03	0.744517
year	-5.822e-04	0.614883
marital_status	-1.332e-01	< 2e-16
education	2.328e-02	7e-06
hh_size	6.620e-03	0.142341
hh_workers	2.315e-02	0.000184

Variable	Estimate	P_value
female	1.157e-01	< 2e-16
age	-3.737e-02	< 2e-16
age_sq	3.569e-04	< 2e-16
imp_hh_rank:post_2001	1.024e-01	0.000893

By applying the first model, we observe that the income rank is 0.31 positively correlated to happiness. P-value of income rank is less than 0.001, we have very strong evidence against the null hypothesis that there is no difference in the model with/without income rank. Besides, we notice that marital status and age are negatively correlated to happiness, while education, number of workers in the household, female, age square, the interaction between income rank and $I2001\check{ }2013$ are positively affecting happiness.

Table 2: Summary statistics in mod2

Variable	Estimate	P_value
(Intercept)	7.593e-01	< 2e-16
imp_hh_rank	3.211e-01	< 2e-16
post_2001	-1.095e-02	0.7446
trend	-3.420e-04	0.8691
marital_status	-1.333e-01	< 2e-16
education	2.330e-02	6.95e-06
hh_size	6.626e-03	0.1420
hh_workers	2.328e-02	0.0002
female	1.157e-01	< 2e-16
age	-3.737e-02	< 2e-16
age_sq	3.569e-04	< 2e-16
year	NA	NA
imp_hh_rank:post_2001	1.093e-01	0.0603
imp_hh_rank:trend	-4.998e-04	0.8891

We add the interaction of income rank with the period of 2001-2013 and the interaction of income rank with the trend. In this case, we find income rank is 0.32 positively significantly associated with happiness.

Table 3: Summary statistics in mod3

Variable	Estimate	P_value
(Intercept)	3.165e+00	0.239750
imp_hh_rank	3.094e-01	< 2e-16
post_2001	2.460e-03	0.933411
pre_2001	3.098e-03	0.914225
marital_status	-1.333e-01	< 2e-16
education	2.334e-02	6.68e-06
hh_size	6.709e-03	0.137173
hh_workers	2.221e-02	0.000387
female	1.155e-01	< 2e-16
age	-3.744e-02	< 2e-16
age_sq	3.576e-04	< 2e-16
year	-1.208e-03	0.372410
imp_hh_rank:post_2001	1.117e-01	0.001514
imp_hh_rank:pre_2001	2.471e-02	0.598603

We keep adding the interaction of income rank with at the period of 1985-2000 (before the tax agency released the income data). We have strong evidence that interaction between income rank and post period is 0.11 positively significantly associated with happiness (p-value is large than 0.001 but less than 0.05).

Table 4: Summary statistics in mod4

Variable	Estimate	P_value
(Intercept)	6.971e-01	0.763210
imp_hh_rank	3.507e-01	< 2e-16
post_2001	3.127e-02	0.225897
higher_internet	1.245e-01	0.000132
marital_status	-1.323e-01	< 2e-16
education	1.752e-02	0.001559
hh_size	2.003e-03	0.670392
hh_workers	2.049e-02	0.001325
female	1.214e-01	< 2e-16
age	-3.665e-02	< 2e-16
age_sq	3.526e-04	< 2e-16
year	1.392e-05	0.990457
imp_hh_rank:higher_internet	-1.280e-01	0.014546
post_2001:higher_internet	-1.736e-01	2.65e-05
imp_hh_rank:post_2001	1.125e-02	0.818059
imp_hh_rank:post_2001:higher_internet	2.670e-01	0.000155

The triple-difference specification model states higher internet is 0.12 significantly increased the happiness, interaction of income rank with post-2001 and higher internet is also raise 0.27 of happiness. In contrast, the interaction of income rank with higher internet, the interaction of post-2001 with higher internet low down happiness gradient.

As we mentioned above, subjective well-being is measured by both happiness and life satisfaction. Hence, we also use life satisfaction as a responding variable and apply the models. The results are included in Appendix. Both of the models support income rank raises life satisfaction (p-values are less than 0.001).

5   Discussion

As GDP grows rapidly in many countries, overall national income raises. Economists argue that whether income can affect individual subjective well-being. The original paper "The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment" analyze the relationship between Norwegian income transparency and individual well-being since the residents are allowed to access the income information online since 2001 (Perez-Truglia 2020). We intend to replicate the paper using the analytic datasets which can be generated by the provided codes. Similar to the original paper, we want to study the relationship between income rank and well-being (both happiness and life satisfaction). We first plot the histogram to show the Google search volumes of the four keywords: Youtube, weather, tax, and tax lists in Norway. We can observe that Norwegians are much more likely to search tax lists, meaning they are interested in the tax search tool. Also, the search of tax lists increases sharply when the agency released the income data. Then we build the models followed by the original paper. To control the time effect and other biases, adding a triple-difference specification and the variable of the pre-experimental year are necessary. Unlike the original paper also compare the results between Norway and Germany, we excluded this part. For some reason, we are unable to access the Germany dataset. So we only conduct the models using Norway dataset.

We find our estimations prove the Easterlin Paradox which was demonstrated by economists Richard Easterlin. It shows that happiness is directly relatives to income within nations, while higher income is associated

with higher happiness. Likewise, the relative income (RI) hypothesis also explains the relationship (Duesenberry 1949). The hypothesis shows that individual compares their income to others, and happiness depends on individual relative income rank (Duesenberry 1949). Many evidence supports the hypothesis. For instance, the researchers estimate the UK household dataset and conduct a longitudinal study (Brown, Gray, and Jennier 2015). They use life satisfaction and General Health Questionnaire (GHQ) as dependent variables to measure utility proxies. They find the RI positively and significantly correlated to the two dependent variables in the spatial (region, occupation, industry) group (Brown, Gray, and Jennier 2015). The results help economists and policy makers realize individual income plays an essential role in individual well-being. It is necessary to for us develop national income and GDP.

In our report, some weaknesses should be improved in the future. First, the survey is sent by mail which asks respondents to complete several questions (Perez-Truglia 2020). But non-respond bias exists. For example, some potential respondents may decide not to complete the survey form because they think they are not very happy. Second, we are supposed to set Norway data as a treatment group, Germany data as a control group in the difference in difference models. Then we compare the pre and post-period of individual well-being. It is better to do so since it can reduce the time effect and bias, also help us predict the actual policy effect. Third, there are many specifications besides higher internet (the internet access at home) that may affect subjective well-being. Thus, we should consider including more specifications in future studies.

Appendix

Table 5: Summary statistics in mod5

Variable	Estimate	P_value
(Intercept)	−1.208e+01	9.56e−05
imp_hh_rank	6.295e−01	< 2e−16
post_2001	2.848e−02	0.4204
year	6.137e−03	7.44e−05
marital_status	−6.958e−02	< 2e−16
education	4.820e−02	2.65e−12
hh_size	−5.579e−03	0.3424
hh_workers	−1.061e−02	0.2493
female	1.531e−01	< 2e−16
age	−3.702e−02	< 2e−16
age_sq	−3.702e−02	< 2e−16
imp_hh_rank:post_2001	1.142e−01	0.0475

Table 6: Summary statistics in mod6

Variable	Estimate	P_value
(Intercept)	−1.336e+01	1.76e−05
imp_hh_rank	5.674e−01	3.00e−10
post_2001	5.142e−02	0.2154
higher_internet	3.540e−02	0.6441
year	6.788e−03	1.31e−05
marital_status	−7.323e−02	< 2e−16
education	4.769e−02	1.69e−11
hh_size	−1.151e−02	0.0595
hh_workers	−1.197e−02	0.2166
female	1.576e−01	< 2e−16
age	−3.690e−02	< 2e−16
age_sq	4.193e−04	< 2e−16
imp_hh_rank:higher_internet	4.116e−02	0.7484
post_2001:higher_internet	−1.081e−01	0.1765
imp_hh_rank:post_2001	5.000e−02	0.5928
imp_hh_rank:post_2001:higher_internet	1.784e−01	0.1910

References

Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, Winston Chang, and Richard Iannone. 2020. *Rmarkdown: Dynamic Documents for R*. https://github.com/rstudio/rmarkdown.

Brown, Sarah, Daniel Gray, and Jennier. 2015. "The Relative Income Hypothesis: A Comparison of Methods." *Economics Letters* 130: 47–50. https://doi.org/10.1016/j.econlet.2015.02.031.

Duesenberry, James. 1949. *Income, Saving, and the Theory of Consumer Behavior.* Harvard University Press.

E.Clark, Andrew, Paul Frijters, and Michael A. Shields. 2008. "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature* 46 (1): 95–144. https://doi.org/10.1257/jel.46.1.95.

Esterlin, Richard A. 2005. "Feeding the Illusion of Growth and Happiness: A Reply to Hagerty and Veen-Hoven." *Social Indicators Research* 74 (3): 429–43. https://doi.org/10.1007/s11205-004-6170-z.

Perez-Truglia, Ricardo. 2020. "The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment." *American Economic Review* 110 (4): 1019–54. https://doi.org/10.1257/aer.20160256.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Evan Miller. 2020. *Haven: Import and Export 'Spss', 'Stata' and 'Sas' Files.* https://CRAN.R-project.org/package=haven.

Xie, Yihui. 2021. *Bookdown: Authoring Books and Technical Documents with R Markdown.* https://github.com/rstudio/bookdown.