

Replication of a Research Claim from Fitzgerald et al. (2018),  
from *Social Forces*

Replication Team: Kent Jason Cheng and Daniel Mallinson

Research Scientist: Sam Field

Action Editor: Annette Brown

Independent Reviewers

(add name below when you initiate review, comment “DONE” on your name when you finish):

Reviewer #1: [Onurcan Yilmaz]

Reviewer #2: Bert Bakker

Reviewer #3: [Anna Szabelska]

Review Period: October 9 - October 14

View-only links to: [Original Paper](#), [Replication Data](#), [Replication Analysis](#)

Privacy Statement: Other teams are making predictions about the outcomes of many different studies, not knowing which studies have been selected for replication. As a consequence, the success of this project requires full confidentiality of this peer review process. This includes privacy about which studies have been selected for replication and all aspects of the discussion about these replication designs.

## Instructions for Data Analysts

The preregistration for this replication study was started by a separate team of researchers who were responsible for identifying data sources and constructing them into a replication dataset(s) for your use in the analysis. They have completed sections 1-13 of the preregistration below, and included additional materials in the OSF project that document how the dataset was constructed.

In cases where all of the underlying data sources were able to be freely shared and posted, the constructed dataset(s) have been posted to the OSF as well, which you are free to use in designing the analysis plan (see below for details). In cases where some or all of the data sources could *not* be freely shared or posted, the replication dataset(s) are not provided on the OSF. Rather, you will need to follow the instructions and code to first reconstruct the datasets, and then proceed with your work. In such cases, the team responsible for creating the dataset(s) has provided summary statistics in the OSF that correspond to the constructed datasets, so you can verify that the datasets you create match what they intended.

You'll be responsible for filling out sections 16-25 of the preregistration below. Before you do so, **please review the original study, sections 1-15 of the preregistration, and the materials provided on the OSF**, so that you are familiar with all of the decisions that have been made to date. In many cases, the 'data preparer' will have left you instructions and suggestions on how the provided data can be used in the analysis, as well as idiosyncrasies and discrepancies in the data that you should be aware of. The data preparers have tried to be thorough in including all variables that you might need, but please keep in mind the following:

- Some of the variables included in the constructed dataset(s) may not be needed in the final analysis, so please do not feel the need to necessarily use all of the provided variables.
- Some of the variables needed might have mistakenly been excluded from the constructed datasets. If you find that this is the case, please let [Andrew](#) or [Anna](#) know, and they will work with you to supplement the datasets as needed.

For these secondary data replications, we would like the analysis plan to be completed before the preregistration goes through review, so that after review, the only remaining steps are registration and running the analysis code on the full datasets. To facilitate that, we are asking that you include in section 19 a link to the code you will use that takes the constructed dataset(s) provided to you and produces the focal analysis (including all of the cleaning, merging, and transforming required). **When developing your analysis plan and code, please randomly sample 5% of the data for use in your work and demonstrate that the focal analysis produces sensible results using just that random sample by providing a screenshot of the output (see section 19 for details). Do not use the rest of the data until after your study is registered and it is time to run the final analysis.** In section 19, you will find a statement that we are asking you to bold that confirms you've only used 5% of the data when developing and testing your code. If this approach will not work for any reason, please let [Andrew](#) or [Anna](#) know and disclose deviations from this plan somewhere in the preregistration.

- In cases where we are providing you a complete dataset, you can just sample out 5% of the observations and hold the rest out until you are ready to perform the final analysis.
- In cases where we are providing you multiple datasets that need to be combined prior to analysis, please sample out 5% of the observations in whatever way is most sensible.
  - For example, in cases where each dataset contains complete observations on its own (a typical 'row bind' situation), it makes the most sense to sample out 5% of each dataset separately and then combine them together to develop and test your code.

- In cases where datasets need to be merged in order to create complete observations (a typical 'column bind' situation), it makes the most sense to merge the separate datasets into a full dataset first, and then sample out the 5% before proceeding with the rest of the analysis code.
- We leave the decision on how to sample out the random subset of data to you, so long as (a) you are not performing any analyses on the complete dataset until after your study is registered and (b) whatever decision you make is documented in the preregistration.

Finally, in cases where the replication data combines observations from the original study with observations that were not used in the original study (what we are calling 'hybrid replications'), please perform up to three analyses (details immediately below). This will likely require you to subset your data, based on the description of the original analysis provided in the study.

- When the 'new' data alone can clear the minimum power threshold, please perform one analysis that relies only on the 'new data' (the focal analysis), one analysis that relies on all available data, and a third analysis that relies only on the original data. Please make sure all three analyses are documented (with code) in section 19 below.
- When the 'new' data alone *cannot* clear the minimum power threshold, please perform one analysis that combines all available data, and a second that only uses the old data. Please make sure both analyses are documented (with code) in section 19 below.

**Please contact [Andrew](#) or [Anna](#) if you have any questions. After you've completed the remaining sections of the preregistration and uploaded all the necessary materials to the OSF, please contact [the SCORE coordinators](#) regarding next steps.**

**Preregistration of Fitzgerald\_SocialForces\_2018\_4q0L**

**Existing Data Replication**

## Study Information

### 1. Title (provided by SCORE)

**RR TEAM INSTRUCTIONS:** *This has been determined by SCORE.*

Replication of a research claim from Fitzgerald et al. (2018) in *Social Forces*.

### 2. Authors and affiliations

**RR TEAM INSTRUCTIONS:** *Fill in the names and affiliations of your team below.*

Kent Jason Cheng, Data Finder<sup>1</sup>

Daniel J. Mallinson<sup>2</sup>

1 Department of Social Science, Maxwell School of Citizenship and Public Affairs, Syracuse University, Syracuse, New York, USA

2 School of Public Affairs, Penn State Harrisburg, Middletown, Pennsylvania, USA

### 3. Description of study (provided by SCORE)

**RR TEAM INSTRUCTIONS:** *This description has been provided by SCORE. Please review and make a SCORE project coordinator aware of any edits, additions, and corrections you would suggest to the paragraph. You are free to add additional descriptions of your project in a separate paragraph.*

The claim selected from Fitzgerald et al. (2018) is that working time is positively associated with higher state-level carbon emissions. This reflects the following statement from the paper's abstract: "Our findings suggest that over the 2007–2013 period, state-level carbon emissions and average working hours have a strong, positive relationship, which holds across a variety of model estimation techniques and net of various political, economic, and demographic drivers of emissions." The authors estimate Prais-Winsten models with panel corrected standard errors (PCSEs) for the fixed effects models. They estimate two-way fixed effects models by including both state-specific and year-specific intercepts. They also correct for first-order auto-correlation (i.e., AR(1) correction) within panels and treat that AR(1) process as common to all panels. The specification of the model containing the selected test result can be gleaned from Table 4, Model 1: Scale(FE). The selected test involves the location of the estimated coefficient in the "Working hours" term. In model 1 (Table 4), the authors find that the scale effect of working hours on carbon emissions is positive and significant. Specifically, in model 1, they find that,

over time, a 1 percent increase in average working hours per worker is associated with a 0.668 per-cent increase in emissions, holding all else constant ( $b = .668$ ,  $SE = .179$ ,  $p < .001$ ).

#### 4. Hypotheses (provided by SCORE with possible Data Analyst additions)

**RR TEAM INSTRUCTIONS:** *The focal test for SCORE is indicated as  $H^*$ . If you will test additional hypotheses (or use alternate analyses) that help you to evaluate the claim your replication/reproduction is testing, number them  $H_1$ ,  $H_2$ ,  $H_3$  etc. (You can place  $H^*$  in the list wherever makes sense). Please make sure that any additional hypotheses are logical deductions/operationalizations of the selected SCORE claim or are necessary to properly interpret the focal  $H^*$  hypothesis. Research that is outside this scope should be described in a separate preregistration.*

*Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):*

- Are the listed hypotheses specific, concise, clearly testable, and specified at the level of operationalized variables?
- Are hypotheses identified as directional or non-directional, and, if applicable, have the direction of hypotheses been stated? (Example: “Customers’ mean choice satisfaction will be higher in the CvSS architecture condition than in the standard attribute-by-attribute architecture condition.”)
- Does the list of hypotheses/tests indicate whether additional hypotheses are taken from the original study or modified/added by the team?

$H^*$ : Average working hours per worker in a state will be positively associated with carbon emissions.

#### Design Plan

#### 5. Study type

**NOTE:** *The study type selected should be based on the data collected for the replication, and not necessarily the data used in the original study.*

- Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.
- **Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.**
- Meta-Analysis - A systematic review of published studies.
- Other

## 6. Blinding

**RR TEAM INSTRUCTIONS:** Select any/all of the below that apply for your study by bolding them. You will give a longer description in the next question.

- **No blinding is involved in this study.**
- For studies that involve human subjects, they will not know the treatment group to which they have been assigned.
- Personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments. (Commonly known as “double blind”)
- Personnel who analyze the data collected from the study are not aware of the treatment applied to any given group.

### [QUESTION 6 - BOLD YOUR RESPONSE ABOVE]

## 7. Blinding

**RR TEAM INSTRUCTIONS:** Since all existing data replications are based on data that has already been collected, in most cases it will not be necessary to comment on participant blinding. In the rare instance when an existing experiment is being re-analyzed for an existing data replication and blinding is a relevant consideration, please provide below any details regarding blinding that are important for a reviewer to be aware of.

Blinding is not applicable to this study as data is at the US state level.

## 8. Study Design

**RR TEAM INSTRUCTIONS:** Please describe how data was collected in the original study and how it compares to the data that was selected for the replication attempt. Explain why the data selected for the replication study is suitable for a replication and if any substantial deviations exist between the two.

If the data used in the replication combines observations from the original study with new observations (e.g. if the data selected for the replication attempt comes from the same longitudinal survey as the original study), describe how ‘original’ and ‘new’ observations relate to each other and an estimate for what proportion of the final dataset’s observations will be comprised of original vs. new observations.

Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):

- Does the preregistration specify the unit of analysis?

- Does the preregistration provide sufficient detail about how the data selected for the replication attempt deviates from or is congruent with the data employed in the original study?
- Does the preregistration describe whether and how ‘original’ and ‘new observations’ are combined together for the replication dataset?

The study by Fitzgerald et al 2018 aimed to test the relationship between working hours and carbon dioxide emissions in the US using state-level data from 2007-2013. The authors used state-level carbon dioxide emissions from [fossil fuel combustion](#), measured in million metric tons (MMTCO<sub>2</sub>) from the US Environmental Protection Agency (EPA) which can be downloaded from this [link](#).

The independent variable of main interest - average weekly working hours per worker was found from the original source: the US Bureau of Labor Statistics (BLS) Current Employment Statistics (CES) database (2016). However, the BLS website only had data from 2008 to 2019 posted as of the time when I started doing the replication. An alternative is to use the Federal Reserve Bank of St. Louis’s [FRED](#) data service which has the weekly working hours per worker data from 2007 to 2019.

Majority of the other control variables available for the replication are the same with the original. Labor productivity was found in the [BLS](#) while population and state energy production were found in the [EIA](#). To arrive at the manufacturing sector as a percent to GDP, the manufacturing GDP in current dollars should be divided by the GDP in current dollars. Manufacturing as a percent to GDP can be obtained by dividing manufacturing sector GDP to current GDP found in the [Bureau of Economic Analysis \(BEA\) website](#). The [BEA](#) also provided employment rate.

For GDP per capita however, the authors used GDP chained to 2007 dollars but the GDP data available now are chained to 2012 dollars. To get GDP per capita chained to 2012 dollars, GDP chained to 2012 dollars must be divided by total population.

Another difference between the original and the replication lies with the control variables coming from the US Census Bureau’s American Community Survey (ACS), specifically the working age population and average household size. State-level data can be obtained from this [link](#) that allows the analyst to do an advanced search for what is needed. However, the [link](#) only displays ACS 2010-2018. Older ACS can be accessed by either downloading [summary file data](#) from each state, each year, and each survey subtopic (e.g. demographics, employment). However, using the summary file data would require the use of SAS. Since I cannot use SAS, I decided to find other alternatives. I was able to download working age population at the US Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research ([CDC Wonder](#)). On the other hand, I was able to compute the average household size per state in [IPUMS](#). Both CDC Wonder and IPUMS data have used ACS for their numbers so technically the replication exercise still used the original’s source.

Considering the varied time period availability of each variable, the replication dataset can range from 2007-2016 or 10 years. The original study used data from 2007-2013 or 7 years. Save for the dependent variable and the GDP used, the replication dataset extends the original study by 3 years and the original study’s data is 70% of the replication data.

## 9. Randomization (free response)

**RR TEAM INSTRUCTIONS:** *If the variables used for this replication attempt were randomized, state how they were randomized, and at what level.*

Randomization is not applicable in this study as it included 50 US states (excluding District of Columbia).

## Sampling Plan

*This section describes how the data sources for the replication were selected, how they were prepared into a replication dataset, and the number of observations that will be analyzed from these data. Please keep in mind that the data described in this section are the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.*

## 10. Existing data (multiple choice question, provided by SCORE)

- 1.1.1. Registration prior to creation of data
- 1.1.2. Registration prior to any human observation of the data
- 1.1.3. Registration prior to accessing the data
- 1.1.4. Registration prior to analysis of the data**
- 1.1.5. Registration following analysis of the data

## 11. Explanation of existing data

**NOTE:** *For replications that rely on existing data sources, this question refers to the data that will be used for the replication analysis (i.e. the final replication dataset), and not (a) the data from the original study or (b) the data sources accessed to construct the replication dataset. Since no new data will be created for ‘existing data replications,’ 1.1.1 should never be selected. Since all analyses will occur after registration, 1.1.5 should also never be selected.*

Data have been downloaded and merged from the various sources described above. The replication data extends the original data by 3 years. The dependent variable selected for the replication is slightly different from the original (state energy-related carbon dioxide emissions versus state energy-related carbon dioxide emissions from fossil fuel combustion, respectively). Majority of the other control variables available for the replication are the same with the original except for those that involve GDP (GDP per capita, Labor productivity). The original was chained to 2007 values while the replication GDP was chained to 2012 values. The authors say

that they got their ACS-sourced control variables (working age population, average household size) from the Census Bureau website itself but given the constraints mentioned above, the replication will use the working age population downloaded from the CDC Wonder database and average household size from IPUMS. Both the CDC Wonder and IPUMS data came from ACS so technically the main difference between the original and replication for these variables lies in the source of data.

## 12. Data collection procedures

**RR TEAM INSTRUCTIONS:** *Please describe the process for constructing the replication dataset in as much detail as you can. The sections below should be used to provide the following information:*

- *Which variables are needed from the original study to perform a good-faith, high-quality replication.*
- *Which data sources were used, why they were selected, any deviations between the original study design and the replication study design that these selections present, and the procedures used to access the data.*
- *Which of the variables from the original study are available in the replication data sources, including relevant details about each measure.*
- *The procedure for creating the replication dataset, in both narrative and script form.*
- *A data dictionary that documents each variable included in the replication dataset.*

*In the sections below, please provide links to the original materials whenever possible -- including descriptions of the original datasets and corresponding codebooks. If materials can be shared on the OSF, please do so, and provide view-only links to those materials.*

*Specific points to keep in mind for reviewers:*

- *Does the preregistration describe which data sources were selected for the replication study and why each is suitable?*
- *Does the preregistration make clear how the data sources were used to construct the replication dataset?*

### (a) Data Needed

**RR TEAM INSTRUCTIONS:** *List below the datasets and variables the original author used to analyze the focal claim. Include details regarding the sample size, waves or years used, and other details pertinent to finding an existing dataset for replication. Please include page numbers when excerpting from the original article. If possible, categorize the list of variables as one of the following: dependent variable, focal independent variable, control variable, or sample parameters/clustering variable. Finally, include the sample size of the original study's focal analysis, if it is available.*

The authors compiled the variables from multiple sources.

### **Dependent Variable(s)**

"state-level carbon dioxide emissions from fossil fuel combustion, measured in million metric tons (MMTCO<sub>2</sub>)" (p 9)

### **Focal Independent Variable(s)**

Average weekly working hours per worker

"We gathered these data from the US Bureau of Labor Statistics (BLS) Current Employment Statistics (CES) database (2016). Each month, the CES surveys approximately 146,000 businesses and government agencies, covering approximately 623,000 individual worksites. The sample of businesses for each month comes from the BLS Longitudinal Database of employer records, which contains data on approximately 9.3 million businesses across the United States (BLS 2016). The working hours data cover all nonfarm private employees, but exclude public employees." (p 10)

### **Control Variable(s)**

Labor productivity

- "Labor productivity is measured as GDP per hour of work." (p 10)
- "The data on labor productivity and employed population were also gathered from the BLS (2016)." (p 10)

% of Employed population

- "Employed population percentage is measured as the employed population of a state divided by its total population." (p 10)
- "The data on labor productivity and employed population were also gathered from the BLS (2016)." (p 10)

Total population size, manufacturing as a % to GDP, states' energy production

- "Other control variables include total population size (US Census Bureau 2016), manufacturing as a percentage of GDP (US Department of Commerce Bureau of Economic Analysis 2016), and a state's energy production, which contributes to carbon emissions. The state energy production data were collected from the EIA's State Energy Database System (2016)." (p 10)

% of Working Age Population and average household size

- "We also control for the working-age population percentage (measured as the percentage of the population aged 15–64) and average household size. These data were obtained from the US Census Bureau's American Community Survey (2017)." (p 10)

Real GDP

- "State-level GDP per capita data (in chained 2007 dollars) were taken from the US Department of Commerce Bureau of Economic Analysis (2016) database." (p 10)

### **Sample Parameters**

The time period for the analyses range from 2007 and 2013 and they include all 50 US states (excludes District of Columbia).

(b) Data Access

**RR TEAM INSTRUCTIONS:** *Describe below the data sources that will provide the replication variables. Include information such as the name of the data source (e.g., Indonesian Family Life Survey), the description and link of the data source, and the waves needed to create a final replication dataset.*

*Also describe the process for accessing the data sources that will be used to create the final replication dataset; specify how long long it took for the registration to be approved and what information was required (e.g., writeup of the purpose of the project, email address from an IPCSR institution, etc.); and verify that the data can be opened as expected. If applicable, provide a link to the page where you registered to access the data.*

*Describe in detail any restrictions on data access and data-sharing, as well as any additional terms of data use that will be relevant for the replication study and final report (e.g. citations that will need to be made). If you were able to access the data because of special permissions that you have, but that you expect other researchers might not have, please document those as well.*

As mentioned, the variables are downloaded from various sources as either xlsx or csv like the EPA (see following [link](#) and click on “co2ffc\_2017 (XLXS)” to download state-level carbon dioxide emissions from fossil fuel combustion, measured in million metric tons (MMTCO2)), the EIA (see [GDP chained to 2012 dollars, population, state energy production](#)), BLS (see links for [labor productivity](#)), BEA (click on the [SAGDP tables: Annual GDP by State](#) link for variables to be used for computing manufacturing as a % to GDP and for employed population, click on this [link](#) then click Annual personal income and employment by state>total full-time and part-time wage and salary employment by industry (SAEMP27) > NAICS(1998-forward) > filter to all states and "total employment (number of jobs)", and CDC Wonder (do an advanced search on this [link](#)).

I could not find any suggested citation for these sources except for CDC Wonder which provides it along with the downloaded data: United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Bridged-Race Population Estimates, United States July 1st resident population by state, county, age, sex, bridged-race, and Hispanic origin. Compiled from 1990-1999 bridged-race intercensal population estimates (released by NCHS on 7/26/2004); revised bridged-race 2000-2009 intercensal population estimates (released by NCHS on 10/26/2012); and bridged-race Vintage 2018 (2010-2018) postcensal population estimates (released by NCHS on 6/25/2019). Available on CDC WONDER Online Database. Accessed at <http://wonder.cdc.gov/bridged-race-v2018.html> on Jun 30, 2020 4:01:21 PM. The EPA website did not specifically mention any suggested citation but it provided this reference to the data:

EPA 2019. Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990-2017. U.S. Environmental Protection Agency, Washington, D.C. April 2019. EPA 430-R-18-003

On the other hand, annual average weekly hours of all employees in the private per state were obtained from the [FRED](#) website. The FRED website organized their files on a per state basis and here are the links: [Alabama](#), [Alaska](#), [Arizona](#), [Arkansas](#), [California](#), [Colorado](#), [Connecticut](#), [Delaware](#), [Florida](#), [Georgia](#), [Hawaii](#), [Idaho](#), [Illinois](#), [Indiana](#), [Iowa](#), [Kansas](#), [Kentucky](#), [Louisiana](#), [Maine](#), [Maryland](#), [Massachusetts](#), [Michigan](#), [Minnesota](#), [Mississippi](#), [Missouri](#), [Montana](#), [Nebraska](#), [Nevada](#), [New Hampshire](#), [New Jersey](#), [New Mexico](#), [New York](#), [North Carolina](#), [North Dakota](#), [Ohio](#), [Oklahoma](#), [Oregon](#), [Pennsylvania](#), [Rhode Island](#), [South Carolina](#), [South Dakota](#), [Tennessee](#), [Texas](#), [Utah](#), [Vermont](#), [Virginia](#), [Washington](#), [West Virginia](#), [Wisconsin](#), [Wyoming](#). The suggested citation for Alabama is: Federal Reserve Bank of St. Louis and U.S. Bureau of Labor Statistics, Average Weekly Hours of All Employees: Total Private in [Alabama \[SMU0100000500000002A\]](#), retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/SMU0100000500000002A>, July 2, 2020.

The above mentioned data can be shared. The only piece of the replication data that cannot be shared is the proxy for average household size per state downloaded from [IPUMS](#) according to this link. Registration is required but access is granted almost immediately after signing up. The main steps in IPUMS is to first select the [sample concerned](#) (tick the box that corresponds to the 2007 to 2016 annual ACS) then click submit sample selection. After which, click search and search for the variable named NUMPREC. According to this [link](#), IPUMS data is to be cited as follows: Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. <https://doi.org/10.18128/D010.V10.0>.

(c) Variable Availability

**RR TEAM INSTRUCTIONS:** For each variable required for the replication analysis (listed above), describe the variables from the replication data that can be used to measure it (including which data files or sources each measure is found in), any notes a data analyst should consider when using the measure in a replication analysis, and any important differences between the original variable and the proposed replication variable.

If there are multiple variables in the replication data that correspond to a required variable (e.g. two different measures of education in the replication data), include all of those options below. If a variable from the original study cannot be measured using the replication data, please make that clear as well. Finally, include a description of the identifiers used to merge multiple datasets, if applicable.

Note that I created the variable names as most raw files did not have variable names that fit typical dataset naming conventions. The analyst must also take the logs of these variables as “All non-binary variables are transformed into logarithmic form, an established approach in research on the human drivers of anthropogenic emissions and related outcomes” (p 9)

Variable name: epa

- File name: co2ffc\_2017\_2.xlsx
- Folder: Raw files\Addendum
- Description: State-level carbon dioxide emissions from fossil fuel combustion, measured in million metric tons (MMTCO<sub>2</sub>)
- Additional notes: This is the actual dependent variable used by Fitzgerald et al which was found in October 13, 2020. There is no longer a need to use the substitute variables namely carbon\_unadj and carbon\_adj.

### **Focal Independent Variable(s)**

Average weekly working hours per worker

Variable name: wrkhrs

- File name (each state has its own file from [FRED](#), see specific links above):  
SMU01000000500000002A.xlsx, SMU02000000500000002A.xlsx,  
SMU04000000500000002A.xlsx, SMU05000000500000002A.xlsx,  
SMU06000000500000002A.xlsx, SMU08000000500000002A.xlsx,  
SMU09000000500000002A.xlsx, SMU10000000500000002A.xlsx,  
SMU11000000500000002A.xlsx, SMU12000000500000002A.xlsx,  
SMU13000000500000002A.xlsx, SMU15000000500000002A.xlsx,  
SMU16000000500000002A.xlsx, SMU17000000500000002A.xlsx,  
SMU18000000500000002A.xlsx, SMU19000000500000002A.xlsx,  
SMU20000000500000002A.xlsx, SMU21000000500000002A.xlsx,  
SMU22000000500000002A.xlsx, SMU23000000500000002A.xlsx,  
SMU24000000500000002A.xlsx, SMU25000000500000002A.xlsx,  
SMU26000000500000002A.xlsx, SMU27000000500000002A.xlsx,  
SMU28000000500000002A.xlsx, SMU29000000500000002A.xlsx,  
SMU30000000500000002A.xlsx, SMU31000000500000002A.xlsx,  
SMU32000000500000002A.xlsx, SMU33000000500000002A.xlsx,  
SMU34000000500000002A.xlsx, SMU35000000500000002A.xlsx,  
SMU36000000500000002A.xlsx, SMU37000000500000002A.xlsx,  
SMU38000000500000002A.xlsx, SMU39000000500000002A.xlsx,  
SMU40000000500000002A.xlsx, SMU41000000500000002A.xlsx,  
SMU42000000500000002A.xlsx, SMU44000000500000002A.xlsx,  
SMU45000000500000002A.xlsx, SMU46000000500000002A.xlsx,  
SMU47000000500000002A.xlsx, SMU48000000500000002A.xlsx,  
SMU49000000500000002A.xlsx, SMU50000000500000002A.xlsx,  
SMU51000000500000002A.xlsx, SMU53000000500000002A.xlsx,  
SMU54000000500000002A.xlsx, SMU55000000500000002A.xlsx,  
SMU56000000500000002A.xlsx
- Folder: Raw files\Work Hours

- Description: Average weekly hours of all employees: total private by state, annual, not adjusted for seasonality
- Additional notes:

### **Control Variable(s)**

Variable name: laborprod

- File name: lpc-by-state-and-region.xlsx
- Folder: Raw files
- Description: Labor productivity (GDP per hours worked)
- Additional notes: The original data was chained to 2007 dollars but the replication data is chained to 2012 values.

Variable name: pop

- File name: use\_pop\_gdp.xlsx
- Folder: Raw files
- Description: Total population size (Resident population including Armed Forces, thousand)
- Additional notes:

Variable name: rgdp

- File name: use\_pop\_gdp.xlsx
- Folder: Raw files
- Description: Real Gross Domestic Product (GDP), million chained (2012) dollars
- Additional notes: The original data was chained to 2007 dollars but the replication data is chained to 2012 values. Divide by population (pop) to get real GDP per capita.

Variable name: manuf

- File name: SAGDP2N\_\_ALL\_\_AREAS\_1997\_2019.xlsx
- Folder: Raw files\SAGDP.zip
- Description: GDP from manufacturing by state in millions of current dollars
- Additional notes: This must be divided by current GDP (variable name: gdp) to get manufacturing to GDP.

Variable name: gdp

- File name: SAGDP1\_\_ALL\_\_AREAS\_1997\_2019.xlsx
- Folder: Raw files\SAGDP.zip
- Description: GDP by state in millions of current dollars
- Additional notes: Use this as denominator for manufacturing to GDP. Do not use rgdp or real GDP since manufacturing (variable name: manuf) is in current dollars.

Variable name: energy

- File name: prod\_btu\_re\_te.xlsx
- Folder: Raw files
- Description: Total energy production, billion Btu
- Additional notes:

Variable name: workpop

- File name: Bridged-Race Population Estimates 1990-2018.txt
- Folder: Raw files
- Description: Populations aged 15 to 64
- Additional notes: Divide by total population to get the percentage.

Variable name: emppop

- File name: download.xlsx
- Folder: Raw files
- Description: Total employment (Number of jobs)
- Additional notes: Divide by total population to get the percentage.

Variable name: NUMPREC

- File name: usa\_00006.dat
- Folder: N/A (Cannot be shared)
- Description: (from IPUMS website) “NUMPREC reports the number of person records that are included in the sampled unit. These person records all have the same serial number (SERIAL) as the household record. The information contained in the household record usually applies to all these persons.”
- Additional notes: This was downloaded from IPUMS. The analyst must process the zip file according to IPUMS’ instructions, and then compute for the state-level average household size. To adjust the mean estimates, use PERWT and STRATA. “[PERWT](#) indicates how many persons in the U.S. population are represented by a given person in an IPUMS sample. It is generally a good idea to use PERWT when conducting a person-level analysis of any IPUMS sample.” On the other hand, “[STRATA](#) is designed for use with CLUSTER in Taylor series linear approximation for correction of complex sample design characteristics.

Variable name: hsize

- File name: N/A
- Folder: N/A (Cannot be shared)
- Description: This was computed by taking the survey weight adjusted state-level mean of NUMPREC for each year.
- Additional notes: See description of NUMPREC, PERWT, and STRATA as mentioned above.

## Identifiers

Variable name: State

- Description: States’ name in string format
- Additional notes:

Variable name: state\_id\_no

- Description: States’ numerical identifier generated from variable name: states
- Additional notes: This is generated such that the numbering of the states are alphabetical (e.g. Alabama has a state\_id\_no of 1, Alaska has 2, so forth); The

alphabetical arrangement does not include states that are not in the sample (i.e. DC, or other US territories)

Variable name: year

- Description: Time in calendar years
- Additional notes:

#### (d) Data Creation

**RR TEAM INSTRUCTIONS:** Create a dataset using the data sources and variables listed above. Provide a detailed narrative describing how the various datasets were cleaned and merged into a final replication dataset. Provide a view-only link to a clearly commented script on the OSF that produces the replication data as described in the narrative. Our preference is that this be either an R script or a script from another language that similarly allows for open and reproducible analyses. Please let the SCORE team know if this is not possible.

The data from the replication sources should be preserved in as ‘raw’ a form as possible, in order to give the data analyst the most latitude to clean the variables as they see fit. Variables from the original source should be preserved in their original form (e.g. do not recode values of 99 to NA). New variables should only be created when they’re needed to complete the merge or combine the datasets; in those cases, please preserve a version of the original, unaltered variable in the new dataset.

When combining multiple datasets by binding rows, please be sure that the data type and measurement units are equivalent across each dataset. If there is a discrepancy in how a variable is measured across datasets, rename the variable in each dataset to indicate the original dataset, and then carefully document the resulting measures below and in the data dictionary. [See here for an example](#) of how this should work.

Please also use this section to describe:

- Any deviations between the original study design and the replication design that would result from using this replication dataset.
- Any notes about using these variables that you would like to pass along to the data analyst.

I used STATA in the creation of the replication data. I divided the work into three do files with file names that are self-explanatory namely:

- [1] Exporting data to dta format.do
- [2] Processing IPUMS data.do
- [3] Merging all data to dta format and reshaping to long format.do

These do files have to be run in the order at which they appear. Here is the [view-only link](#) to the script files. As mentioned, the majority of the data collected are in the either xlsx or csv or txt file

so they had to be exported to individual dta files in wide format first. Note that in the script posted, \$ should be replaced with the file destination.

More details about the IPUMS-related do file [2] Processing IPUMS data.do can be found through this [readme](#) file. Results of this do file can be cross-checked with the [descriptive statistics](#) that I generated.

#### (e) Data Dictionary

**RR TEAM INSTRUCTIONS:** Create [a data dictionary](#) following [this template](#). Provide below a view-only link to the completed data dictionary included in the OSF project. If the Data Analyst will need to create new variables using the variables in the final replication dataset (e.g. recoding the provided education variable to be in a better format for analysis), please document below your recommendation on how the analyst should do so. Please also document any additional notes regarding the variables in the dataset that do not fit within the provided data dictionary template or the other sections above.

Here is the view-only [data dictionary](#).

### 13. Sample size

**RR TEAM INSTRUCTIONS:** Please report below the analytic sample size(s) in the replication dataset, with reference to however many units or levels are in the data. Please report as much information here as will be helpful for the review committee to be aware of, including differences in sample size resulting from various analytic decisions (e.g. listwise deletion vs multiple imputation). Finally, when the replication combines observations from the original study with new observations, please estimate what proportion of the analytic sample's observations will be comprised of original vs. new observations.

The original data ranged from 2007 to 2013 or 7 years for 50 US states which does not include District of Columbia, so 7 years X 50 states = 350 state-years. The replication data extends the original by 3 years or 2007 to 2016 so that the sample becomes 10 years X 50 states = 500 state-years.

---

Required sample size [provided by the SCORE team]: The primary unit of analysis is the state-year. An estimate of the minimum viable sample size for the data analytic replication is: 97. For comparison, the stage1 required sample size would be: 470 and the stage2 sample size would be: 1056.

### 14. Sample size rationale

*Since all existing data replications are based on data that has already been collected, the rationale for each sample size is simply the number of observations that the replication dataset contains. SCORE replications that rely on new data collection require a sample size that achieves 90% power to detect 75% of the original effect size for the first stage of the analysis, and 90% power to detect 50% of the original effect size for the pooled analysis if a significant effect in the predicted direction is not found after the first round. The power analyses below take the sample size(s) reported in Section 13 as fixed, and calculate the power to detect (a) 75% of the original effect size and (b) 50% of the original effect size.*

For data analytic replications in SCORE, three sample sizes are calculated:

- A minimum threshold sample size, defined as the sample size required for 50% power of 100% of the original effect
- A stage 1 sample size, defined as the sample size needed to have 90% power to detect 75% of the original effect
- A stage 2 sample size, defined as the sample size needed to have 90% power to detect 50% of the original effect

Details about how those sample sizes were calculated for this project [are found here](#).

## 15. Stopping rule (provided by SCORE)

Because the replication data contains a combination of observations that were used in the original analysis and more recent observations, the SCORE team recommends that three analyses be performed:

- One analysis that uses only observations that **were not** used in the original study. This will be the focal replication analysis of the replication study.
- One analysis that uses all available observations.
- One analysis that uses only observations that **were** used in the original study. This may or may not qualify as the focal reproduction analysis of the study.

## Variables

**RR TEAM INSTRUCTIONS:** *The preregistration form divides variables across three questions: manipulated variables, measured variables, and indices (i.e. analytic variables derived from raw variables). For existing data replications, only fill out the ‘Measured variables’ and ‘Indices’ sections. Please do not fill out anything in the ‘Manipulated variables’ section.*

*The raw data of any transformed variable (e.g. reaction time → log reaction time) or any created index should be defined in the ‘Measured variables’ section. Details regarding the variable transformation should be specified in the ‘Transformations’ section. Details regarding the creation of an index should be specified in the ‘Indices’ section.*

*Across these questions, you should define all variables that will later be used during your analysis (including data preparation/processing). You can describe all variables in the preregistration and/or summarize and link to a [data dictionary](#) (codebook) in your repository to answer these questions.*

*If you will share data from your replication, this is also the place to state whether any variables will be removed prior to sharing the dataset (e.g. to reduce risk of participant identification or comply with copyright restrictions on scale items.)*

## 16. Manipulated variables

**RR TEAM INSTRUCTIONS:** *Manipulated variables in this preregistration refer specifically to variables that have been randomly assigned in an experiment. The use of data from an experiment should be rare in existing data replications. If your existing data replication relies on experimental data, please document each manipulated variable as a measured variable, and use the codebook to indicate what each level of the variable corresponds to (e.g. participants assigned to the treatment condition = 1; participants assigned to the control condition = 0). The default language in bold below has been copied into all existing data replication preregistrations.*

**N/A -- not documented for existing data replications.**

## 17. Measured variables

**RR TEAM INSTRUCTIONS:** *Please use this section to document each variable that was used in the original study’s analysis and the role it served (e.g. dependent variable, control variable, sample parameter, etc). For each variable, provide the description of the variable offered in the paper and/or codebook of the original study, the variable in the replication dataset that it corresponds to, and explain any deviations between the two. In cases where an equivalent replication variable was not found, explain how, if at all, you expect it will affect the replication attempt. In cases where you are adding a variable that was not present in the original study, please explicitly state that you are doing so, and explain how, if at all, you expect it will affect the replication attempt.*

*Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):*

- Does the preregistration surface all of the variables needed to replicate the focal analysis?

- Are deviations between the original variables and replication variables documented when needed?

#### carbon\_unadj

- Use in analysis: Dependent variable
- Description from the original study: “State-level carbon dioxide emissions from fossil fuel combustion, measures in million metric tons (MMTCO<sub>2</sub>)” (p. 1859)
- Variable used in replication: State-level carbon dioxide emissions, measured in million metric tons (MMTCO<sub>2</sub>) (Unadjusted)
- Deviation: The dependent variable selected for the replication is slightly different from the original (state energy-related carbon dioxide emissions versus state energy-related carbon dioxide emissions from fossil fuel combustion, respectively).

#### wrkhrs

- Use in analysis: Focal independent variable
- Description from the original study: “average weekly working hours per worker” (pg. 1860)
- Variable used in replication: Average weekly hours of all employees: total private by state, annual, not adjusted for seasonality
- Deviation: None

#### emppop\_pct

- Use in analysis: Independent variable
- Description from the original study: “Employed population percentage is measured as the employed population of a state divided by its total population” (pg. 1860)
- Variable used in replication: Total employed population (emppop) divided by total state population (pop), multiplied by 100
- Deviation: None

#### laborprod

- Use in analysis: Independent variable
- Description from the original study: “GDP per hour of work” (pg. 1860)
- Variable used in replication: Labor productivity (GDP per hours worked)
- Deviation: The original data was chained to 2007 dollars but the replication data is chained to 2012 values

#### pop

- Use in analysis: Independent variable
- Description from the original study: “total population size” (pg. 1860)
- Variable used in replication: Total population size in thousands
- Deviation: None

#### manu\_gdp

- Use in analysis: Independent variable

- Description from the original study: “manufacturing as a percentage of GDP” (pg. 1860)
- Variable used in replication: GDP from manufacturing by state in millions of current dollars (manuf) divided by GDP by millions of current dollars (gdp) multiplied by 100
- Deviation: None

energy

- Use in analysis: Independent variable
- Description from the original study: “state’s energy production” (pg. 1860)
- Variable used in replication: Total energy production, billion Btu
- Deviation: None

hhsize

- Use in analysis: Independent variable
- Description from the original study: “average household size” (pg. 1860)
- Variable used in replication: This was computed by taking the survey weight adjusted state-level mean of NUMPREC for each year
- Deviation: None

workpop

- Use in analysis: Independent variable
- Description from the original study: “measured as the percentage of the population aged 15-64” (pg. 1860)
- Variable used in replication: Populations aged 15 to 64 divided by population total multiplied by 100
- Deviation: None

All of the above variables were log transformed (ln), per the original authors’ analysis.

State

- Use in analysis: Intercepts for state
- Description from the original study: “unit-specific intercepts” (pg. 1866)
- Variable used in replication: State dummy variables
- Deviation: None

factor(year)

- Use in analysis: Intercepts for year
- Description from the original study: “year-specific intercepts” (pg. 1866)
- Variable used in replication: Year dummy variables
- Deviation: None

## 18. Indices

**RR TEAM INSTRUCTIONS:** *If any of the measured variables described in Section 17 will be combined into a composite measure (including simply a mean), describe in detail what measures you will use and how they will be combined. Please be sure this preregistration includes a link to a clearly commented script that constructs the index according to the narrative.*

*Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):*

- Does the preregistration specify each of the composite measures (e.g. mean scores, factor scores) that are needed for the focal analysis, and which of the measured variables in Section 17 are used in each one (e.g. the happiness, joy, and satisfaction items will be used to create the ‘positive feelings’ measure)?
- Does the preregistration link to a clearly commented script that constructs the indices according to the narrative description?

Not applicable

## Analysis Plan

### 19. Statistical models

**RR TEAM INSTRUCTIONS:** This section should describe in detail the analysis that will be performed to replicate the focal result. This analysis must align as closely as possible with the original study’s analysis, even if you have identified limitations in the original study. The level of detail should allow anyone to reproduce your analyses from your description below. Examples of what should be specified: the model; each variable; adjustments made to the standard errors and to case weighting; additional analyses that are required to set up the focal analysis; and the software used.

Beyond the replication of the focal analysis from the original study, it is at your discretion to test the claim using other analytic approaches as a check of the robustness of the claim. The original test should be listed first and be clearly distinguished from any other tests. If you are testing additional confirmatory hypotheses, describe them in the same order as you numbered them in the “Hypotheses” section above and make clear reference to the specific hypothesis being tested for each.

Please provide a link to a clearly commented script that performs the analysis described in the narrative provided below. Our preference is that this be either an R script or a script from another language that similarly allows for open and reproducible analyses. Please let the SCORE team know if this is not possible. Please also test that the code runs without error on a random subset of 5% of the replication dataset, and provide verification that the code has produced a sensible result below by providing a screenshot of the output (please upload the screenshot to the OSF as well). Finally, please confirm that you have only developed and tested your analysis plan and code using 5% of the data.

*Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):*

- Does the preregistration specify which statistical model will be used to provide the ‘focal evidence’ for the SCORE test (e.g. a regression coefficient in a larger multiple regression

*model), and does it correspond closely to the model and evidence from the original study?*

- *Does the preregistration describe each variable that will be included in the focal analysis, and what role each variable has (e.g. dependent variable, independent variable)?*
- *Does the preregistration include a detailed specification of the focal analysis, including interactions, lagged terms, controls, etc., in both narrative form and in a clearly commented script?*
- *Does the preregistration verify that the code runs without error on a random subset of the replication dataset?*

This replication focuses on Model 1 in the original paper. The authors estimated Prais-Winsten models with panel corrected standard errors (PCSEs) for the fixed effects models. They estimate two-way fixed effects models by including both state-specific and year-specific intercepts. They also correct for first-order auto-correlation (i.e., AR(1) correction) within panels and treat that AR(1) process as common to all panels.

The replication analysis was conducted in R. A Prais-Winsten model was estimated using the [prais](#) package. The `prais_winsten` function from the package includes an AR(1) correction. All variables were logged, per the original authors' instructions. Fixed effects were included for both state and year. The following model was estimated:

```
carbon_unadj ~ wrkhrs + emppop_pct + laborprod + pop + manu_gdp + energy + hhszie +  
workpop + State + factor(year)
```

~~This statement confirms that only 5% of the data have been randomly sampled in developing the analysis plan and code contained in this preregistration.~~

Due to the number of coefficients in the model and the small overall sample size ( $n = 500$ ), 10% of the data ( $n = 50$ ) were used to test the code, as opposed to the typical 5%. Additionally, because of the panel model specification, 5 states were randomly chosen, as opposed to choosing 50 random observations. Without the full panels, the code could not otherwise be tested on the sample.

Per the stopping rule, three analyses were coded in R. One includes the new and old data, one includes only the original data, and a third includes only the newly collected data. Below, results for the first two models using only a sample of the data are provided. The third analysis, using only the newly collected data, could not be conducted due to the small sample, but the code for all three analyses can be found at <https://osf.io/gxc2r/>. The first model includes the full 50 samples observations. A screenshot of the main results is included. It does not show the effects for all of the state- and year-specific fixed effects, which were not reported in the original paper.

```

Panel Regression with AR(1) Prais-Winsten correction and panel-corrected standard errors

Balanced Panel Design:
  Total obs.:      50 Avg obs. per panel 10
  Number of panels: 5 Max obs. per panel 10
  Number of times: 10 Min obs. per panel 10

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 26.60356  35.47055  0.750 0.459500
wrkhrs       2.27071   0.57214  3.969 0.000457 ***
emppop_pct   0.83419   0.80128  1.041 0.306751
laborprod     0.99525   0.36386  2.735 0.010692 *
pop          1.63378   4.63946  0.352 0.727364
manu_gdp     -0.04031   0.15809 -0.255 0.800590
energy        -0.11305   0.07480 -1.511 0.141919
hhsizer       1.03052   1.05434  0.977 0.336736
workpop       -3.40437   5.00890 -0.680 0.502297
StateGeorgia  1.31807   0.60986  2.161 0.039375 *
StateMissouri 0.16726   0.27216  0.615 0.543798
StateOregon   -1.62092   0.20019 -8.097 8.14e-09 ***
StateWisconsin -0.04885   0.26215 -0.186 0.853524
factor(year)8 -0.02283   0.01490 -1.532 0.136720
factor(year)9 -0.03207   0.05059 -0.634 0.531197

```

The second model is estimated using observations from the original study's time frame, 2007-2013 (n = 35). The results are shown below.

```

Panel Regression with AR(1) Prais-Winsten correction and panel-corrected
standard errors

Balanced Panel Design:
  Total obs.:      35 Avg obs. per panel 7
  Number of panels: 5 Max obs. per panel 7
  Number of times: 7 Min obs. per panel 7

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 200.503749  59.554927   3.367 0.003927 **
wrkhrs       0.029429  0.941463   0.031 0.975450
emppop_pct   -0.227520  1.095796  -0.208 0.838137
laborprod     0.040516  0.393341   0.103 0.919239
pop          13.440640  6.964691   1.930 0.071554 .
manu_gdp      0.055201  0.177977   0.310 0.760446
energy        -0.102186  0.106820  -0.957 0.352991
hhszie        0.583372  1.251500   0.466 0.647400
workpop      -20.586439  7.738344  -2.660 0.017105 *
StateGeorgia  5.558488  1.323534   4.200 0.000679 ***
StateMissouri 1.385566  0.543295   2.550 0.021390 *
StateOregon   -2.629422  0.329649  -7.976 5.76e-07 ***
StateWisconsin 0.951583  0.480972   1.978 0.065358 .
factor(year)8 -0.002274  0.016212  -0.140 0.890180
factor(year)9 -0.045260  0.058566  -0.773 0.450901

```

A third model using only the newly gathered data (2014-2016) could not be estimated because the sample size was too small (n=15). However, code for this analysis is included for when the full data analysis is conducted.

## 20. Transformations

**RR TEAM INSTRUCTIONS:** *This section should describe how any of the measured variables or composite measures mentioned above will be transformed prior to the analyses listed in Section 19. These are adjustments made to variables **after** measurement or measure creation, and might include centering, logging, lagging, rescaling etc. Please provide enough detail such that anyone else could reproduce the transformations based on the description below. Please be sure this preregistration includes a link to a clearly commented script that performs the transformations described in the narrative provided below.*

Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):

- Does the preregistration specify which of the measured variables or composite measures will need to be transformed prior to the focal analysis?
- For each variable needing transformation, does the preregistration adequately describe the transformations, including any centering, logging, lagging, recoding, or implementation of a coding scheme for categorical variables?

- Does the preregistration link to a clearly commented script that performs each transformation?

All variables included in the above analyses were log (ln) transformed prior to analysis, per the original paper.

## 21. Inference criteria

**RR TEAM INSTRUCTIONS:** This section describes the precise criteria that will be used to assess whether the hypotheses listed above were confirmed by the analyses in Section 19. The default language below only applies to the test of the SCORE claim,  $H^*$ . It is at your discretion to describe the inferential criteria you will use for any additional analyses. They need not rely on p-values and/or the same alpha level we have specified for  $H^*$ .

If the additional analyses will use multiple comparisons, the inference criteria is a question with few “wrong” answers. In other words, transparency is more important than any specific method of controlling the false discovery rate or false error rate. One may state an intention to report all tests conducted or one may conduct a specific correction procedure; either strategy is acceptable.

Criteria for a successful replication attempt for the SCORE project is a statistically significant effect (alpha = .05, two tailed) in the same pattern as the original study on the focal hypothesis test ( $H^*$ ). For this study, this criteria is met by a positive correlation between working hours and CO<sub>2</sub> emissions.

## 22. Data exclusion

**RR TEAM INSTRUCTIONS:** The section below should describe the rules you will follow to exclude collected cases from the analyses described in Section 19. Note that this refers to exclusions **after** the creation of the replication dataset; exclusion criteria that prevent a case from entering the replication dataset in the first place should be detailed in the ‘Data Collection Procedure’ section above. Please be as detailed as possible in describing the rules you will follow (e.g. What is the specific definition of outliers you will use? Exactly how many attention checks does a participant need to fail before their removal from the analytic sample?).

Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):

- Does the preregistration comment on whether any cases included in the replication dataset will be excluded prior to data analysis?
- If yes, does the preregistration provided detailed instructions on how the exclusions will be performed (e.g. Is the definition of outlier provided? Is the number of attention checks failed before a participant is excluded specified?)

No exclusions occurred.

## 23. Missing data

**RR TEAM INSTRUCTIONS:** *The section below should describe how missing or incomplete data will be handled. Please be as detailed as possible in describing the exact procedures you will follow (e.g. last value carried forward; mean imputation) and any software required (e.g. We will use Amelia II in R to perform the imputation).*

*Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):*

- *Does the preregistration comment on how missing or incomplete data will be addressed (e.g. casewise removal, missing data imputation)?*
- *If applicable, does the preregistration specify how many missing variables will lead to a case's removal (e.g. If a subject does not complete any of the three indices of tastiness, that subject will not be included in the analysis.)?*
- *If applicable, does the preregistration describe how missing data imputation will be performed, including relevant software?*

Not applicable

## 24. Exploratory analysis (Optional)

**RR TEAM INSTRUCTIONS:** *If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time. If any exploratory analyses involve additions to the data collection procedure beyond what was performed in the original study (e.g. additional items on the survey; running another condition in the experiment), please describe them below.*

Not applicable

## 25. Other

**RR TEAM INSTRUCTIONS:** *This section serves two purposes. First, please use this section to discuss any features of your replication plan that are not discussed elsewhere. Literature cited, disclosures of any related work such as replications or work that uses the same data, plans to make your data and materials public, or other context that will be helpful for future readers would be appropriate here. Second, please also re-surface any major deviations from earlier in the preregistration that you expect a reasonable reviewer could flag for concern. Give a summary of these deviations, focusing on larger changes and any possible challenges for comparing the results of the original and replication study.*

*Specific points to keep in mind (please also consult the [Reviewer Criteria](#)):*

- Does the preregistration reference other sections of the preregistration where substantial deviations from the original study have been described (including deviations due to differences in location or time compared to the original study)?
- Does the preregistration comment on plans to make the data and materials from the replication study public?

## Final review checklist

**REVIEWER INSTRUCTIONS:** For the following questions, reviewers please indicate whether you can ‘sign off’ on the following items by adding a comment. You can update this response as the lab moves through revisions during the review period!

- Included in this pre-registration are specific materials needed to create a replication dataset:
  - Is the final replication dataset that the research team constructed suitable for performing a high-quality, good-faith replication of the focal claim selected from the original study?
  - Is the procedure for constructing the final replication dataset sufficiently documented that an independent researcher could construct the same dataset following the procedures and code they lay out?
- Included with this pre-registration is a narrative description of how the replication dataset will be used to perform the focal replication analysis, as well as the specific analytic scripts/code/syntax that will be used:
  - Is the analysis plan (including code) that's documented in the preregistration consistent with a high-quality, good-faith replication of the focal claim selected from the original study?
  - Has the data analyst demonstrated that the analysis code works as expected on a random 5% of the final replication dataset?
- I have reviewed all sections of this pre-registration, and I believe it represents a good-faith replication attempt of the original focal claim.

**Additionally, please consider the following if the preregistration potentially includes a reproduction focal analysis:**

- The observations used for the reproduction analysis were collected and measured in the same way as the original study.
- The observations used for the reproduction analysis were analyzed in the same way as the original study.
- The data analyst has demonstrated that their analysis code works as expected on a random 5% of the reproduction data.
- I believe this preregistration represents a good-faith reproduction attempt of the original focal claim.