Replication of a Research Claim from Kollmeyer (2009),
from *American Journal of Sociology*

Replication Team: Andrew Soh and Christopher Limnios

Research Scientist: Sam Field

Action Editor: Nick Weller

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

Reviewer #1: [Anna Szabelska]

Reviewer #2: [Onurcan Yilmaz]

Reviewer #3: [NAME]

Review Period: November 2 - November 9

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 Kollmeyer_AmJournSocio_2009_EJpm

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 Kollmeyer (2009) in *American Journal of Sociology*.

2. Authors and affiliations

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

Andrew Soh[1]
Christopher Limnios[2]

1 University of Hawaii at Manoa
2 Providence College

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 for replication from Kollmeyer (2009) is that there is an indirect effect of the level of a northern, economically advanced country's imports from the South on deindustrialization that goes through national affluence. Specifically, North-South trade increases real per capita incomes of Northern countries and, in the process, indirectly promotes deindustrialization by heightening national affluence beyond levels that would prevail in the absence of global trade. The focal claim concerns the association between imports from the South and national affluence. This reflects the following statement from the paper's abstract: "The results indicate that each factor makes significant contributions to deindustrialization, and that global trade exerts both direct and indirect effects on employment patterns in economically advanced countries." The author tests the selected claim using two-way fixed-effects regression models and panel data on 18 Organization for Economic Cooperation and Development (OECD) countries from 1970 to 2003. The specification of the model can be gleaned from

Figure 2, equations 1a - 1c, and Table 2, Model 4. The focal test result concerns the location of the estimated coefficient “Imports from the South” under the heading “Model 4”. The dependent variable is national affluence. The result was a statistically significant estimated coefficient t for “Imports from the South” under the heading “Model 4” (b= .910, SE = .104, 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 H1, H2, H3 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*: Imports from the South will be positively associated with national affluence.

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.*

No blinding was involved to the second data collectors’ knowledge.

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 original study collected data for "18 OECD countries from 1970 to 2003" (p. 1644) using OECD's STAN database. The replication data uses trade data from UN Comtrade from 1950 to 2011 for 33 OECD member countries (i.e., there are originally 37 OECD countries but following the original author, Chile, Colombia, Mexico and Turkey are considered "South" countries and hence, excluded from the list). The rest of the datasets required for replication (i.e., GDP, population, and unemployment) are from the same source used by the original authors. An additional dataset containing the years each OECD member country joined the OECD is taken from the OECD (variable: "year_ratified"). An additional variable has also been created to identify the original 18 countries that the original author used (variable: "oecd18").

The author defines the "Global North" as countries in Europe and North America but excluded Mexico and Turkey from this definition. They also moved Australia, New Zealand, Japan, Israel, and Korea from the "Global South" to the "Global North". I follow exactly what was done by the original author. The OECD member countries on top of the 18 that the original authors used, except for Chile and Colombia, are all European countries. Hence, there was no need to move them anywhere.

The final replication dataset has all 33 OECD countries from 1950-2011. There are 3 possible datasets for the data analyst:

- Drop countries not part of the original 18 countries that the original author used (use the variable “oced18” to do this)
- Drop countries for the years before they became OECD countries (use the variable “year_ratified” to do this)
- Use all 33 countries from 1950-2011

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.*

Variables used in this study were not randomized.

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.*

The datasets from the 33 countries referenced above have been accessed, cleaned, and merged prior to registration. Variables were selected based on their expected relevance to the replication analysis, and the values present in the data (in cases where the codebook was

unclear). None of the variables were selected because of their likelihood (or not) of leading to a confirmatory result.

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.*

Dependent Variable(s)

National Affluence

- "The variable national affluence equals a country's gross domestic product (GDP) divided by its total population, with GDP expressed in U.S. dollars at prices and purchasing parities (PPP) from the year 2000." (p. 1652)
- "Data come from the OECD's (2006a) Annual National Accounts, volume 1: Comparative Tables" (p. 1653)

Focal Independent Variable(s)

Imports / Exports from the South

- "Here I define the South with the OECD's regional classification scheme, which categories all of the world's countries into six geographic regions: Africa, Asia (which includes the Middle East), Central and South America, Europe, North America, and Oceania." (p. 1654)
- "Drawing on this classification scheme, I define the South as Africa, Asia, Central and South America, and Oceania, and I define the North as Europe and North America." (p.1654)
- "I then make adjustments to these broad regional categories by moving Mexico and Turkey (from North America and Europe, respectively) to the South, and by moving Australia and New Zealand (from Oceania) and Israel, Japan, and South Korea (from Asia) to the North." (p. 1654)
- "From here, I include only categories 5–8 from the standard international trade classification (SITC) scheme in the trade figures. This step eliminates services, agricultural products, raw materials, and other nonmanufactured goods from the measurement of North-South trade. To facilitate international comparison, values for imports and exports are expressed as a percentage of GDP for all countries." (p. 1654)
- "Data come from the International Trade by Commodities Database (OECD 2002, 2005 b), which reports the annual monetary value of imports and exports at the national, regional, and global levels." (p. 1654)

Control Variable(s)

Unemployment

- "The main control variable, unemployment, accounts for…the "failure effect"." (p.1655)
- "Data are taken from Labor Force Statistics – Summary Tables (OECD 2006b)." (p. 1655)

Sample Parameters

- No additional sample parameters other than the ones mentioned above.

“To assess the various explanations for deindustrialization against empirical evidence, I assemble a data set comprising repeat observations of 18 OECD countries from 1970 to 2003. The resulting panel contains a maximum of 612 separate observations (n = 18, t = 34), although some observations are missing data for certain explanatory variables.” (p. 1655 - 1656)

The “Description of the analysis” refers to Model 4 of Table 2. The variables listed above are the variables needed to replicate this test. However, for the other models in the paper, the other variables needed are: *unbalanced productivity growth, imports from the North, exports to the North,* and *net outflow of direct investment*. Data for these other variables were not downloaded.

(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.

Data has been downloaded from the same data source as the original authors (https://stats.oecd.org/). The database does not require registration. Trade data, on the other hand, has been downloaded from UN Comtrade (https://comtrade.un.org/).

OECD Data

For OECD Data, the following policy applies (excerpt copied from OECD website: (http://www.oecd.org/termsandconditions/):

“(c) Data

“The OECD makes data (the “Data”) available for use and consultation by the public. Data may be subject to restrictions beyond the scope of these Terms and Conditions, either because specific terms apply to those Data or because third parties may have ownership interests. It is the User’s responsibility to verify, either directly in the metadata or, if available, by clicking on

the icon and then referring to the "source" tab, whether the Data is fully or partially owned by third parties and/or whether additional restrictions may apply, and to contact the owner of the Data before incorporating it in your work in order to secure the necessary permissions. The OECD in no way represents or warrants that it owns or controls all rights in all Data, and the OECD will not be liable to any User for any claims brought against the User by third parties in connection with the use of any Data."

"Permitted use

"Except where additional restrictions apply as stated above, You can extract from, download, copy, adapt, print, distribute, share and embed Data for any purpose, even for commercial use. You must give appropriate credit to the OECD by using the citation associated with the relevant Data, or, if no specific citation is available, You must cite the source information using the following format: OECD (year), (dataset name),(data source) DOI or URL (accessed on (date)). When sharing or licensing work created using the Data, You agree to include the same acknowledgment requirement in any sub-licenses that You grant, along with the requirement that any further sub-licensees do the same."

UN Comtrade Data

For UN Comtrade Data, the following policy applies (excerpt copied from https://comtrade.un.org/db/help/PolicyOnUseAndRedissemination.pdf):

"Permission for re-dissemination and use of data within applications

"xiv. UNSD allows the use of UN COMTRADE data within data extraction and/or visualization/analytical applications, either free-of-charge or with fees. In both cases, written permission from UNSD is necessary and a royalty fee may be applied. In some cases, users of such applications are required to obtain a yearly subscription to UN Comtrade."

"xv. Permission for the re-dissemination of UN COMTRADE data needs to be requested in writing (contact comtrade@un.org); in general, permission will be granted without application of a royalty fee if the amount of data used does not exceed 1,000 records."

"xvi. For publication of UN COMTRADE data in a few tables or graphs in newspaper articles, journals, other magazines or books, it is not necessary to request permission; please refer to the source of the data as "DESA/UNSD, United Nations Comtrade database"."

"xvii. The following rationale is considered regarding the application of a royalty fee:

a. If data being re-disseminated are substantially different from the data provided in the UN COMTRADE database, then a royalty fee shall not apply;

b. If data being re-disseminated are substantially the same as data provided in the UN COMTRADE database, then a royalty fee shall apply where the amount charged is based on:

1) the number of records being re-disseminated or type of application (data visualization/analytics or data extraction oriented);
2) whether data is re-disseminated for a fee or free-of-charge; and,
3) whether access to this data is limited to a specific audience or open to the public in general.

The actual fees are calculated based on the price of a Premium Site License (PSL), which is $6,065 in 2014 (unchanged since 2009).

"xviii. Exceptionally, case by case decisions will be taken for special situations wherein the previously stated guidelines may not easily be applied."

The royalty fee scheme table can be found via this link.

(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.***

National Affluence

- The national affluence measure needs both GDP and total population data

Gross Domestic Product (GDP)

- Country-level data can be downloaded from https://stats.oecd.org/index.aspx?queryid=60702
- Theme: "GDP, US $, constant, prices, constant PPPs, reference year 2015, millions copy"
- Year: "1950 – 2019"
- Saved as "gdp.csv"

Population

- Country-level data can be downloaded from https://data.oecd.org/pop/population.htm
- Perspectives: “Total”, “Millions persons”
- Countries: “OECD (35)” and “COL” on Select Background
- Time: “Yearly” from 1951 – 2018
- Saved as “population.csv”

Imports from / Exports to the South

- Country-level data can be downloaded from https://comtrade.un.org/Data/
- Type of Product: “Goods”
- Frequency: “Annual”
- Classification: “SITC Rev.1”
- Periods: 1965 to 2019 (5 at a time)
- Reporters: Australia, Austria, Belgium, Canada, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom, and United States (5 at a time)
- Partners: “All”
- Trade Flows: “Import” and “Export”
- SITC Rev. 1 Commodity Codes: 5, 6, 7, 8
- Saved as “Name1Name2YY1toYY2.csv”, where “Name1” is the name of the alphabetical first reporter country in the file, “Name2” is the name of the last alphabetical reporter country in the file, “YY1” is the first of 5 years, and “YY2” is the fifth of five years.

These trade files are bilateral trade data (e.g., Import/Export data by Reporter Counter to each country in the world). To create the North/South dataset, import data should be summed up from all countries in the South for each reporter country. Similarly, for each reporter country, export data to all countries in the South should be summed up.

Note: Chile, Colombia, Mexico, and Turkey are OECD countries but are excluded from the list of trade countries above because: (1) Chile and Colombia are not countries in the North as defined by the original author and (2) Mexico and Turkey were moved from the North to the South by the original author.

North / South Countries

- The list of countries per continent was taken from https://www.worldatlas.com/cntycont.htm

- The categorization from World Atlas, however, does not include a separate category for Central America. So I moved Panama, Costa Rica, Nicaragua, Honduras, El Salvador, Guatemala, and Belize from North America to Central America (see R Script).
- By doing so, the data has 6 geographic regions as the original dataset: Africa, Asia, Central and South America, Europe, North America, and Oceania
- Saved as “continent.csv”

Unemployment

- Country-level data can be downloaded from https://data.oecd.org/unemp/unemployment-rate.htm
- Perspectives: “Total” and “% of labour force”
- Countries / Select Background: “OECD (35)” and “COL”
- Time: “yearly” and “1953 – 2019”
- Saved as “unemployment.csv”

OECD Countries

- The list of OECD member countries can be found at https://www.oecd.org/about/document/list-oecd-member-countries.htm
- According to the website, “on 14 December 1960, 20 countries originally signed the Convention on the Organisation for Economic Co-operation and Development. Since then, 17 countries have become members of the Organisation”.
- These countries are: Australia, Austria, Belgium, Canada, Chile, Colombia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States.
- This file is saved as “oecdratification.csv”

(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.*

- *If the data can be freely shared and posted to OSF, please post it in your OSF project and provide a link to the completed dataset below.*
- *If any part of the dataset cannot be shared between researchers or posted to the OSF, please leave the final dataset off the OSF. Instead, include either below or in your script*

(commented out at the bottom) two pieces of information that will help an independent team verify they have created the dataset according to your instructions:

- *The dimensions of the final dataset(s) you've created (# of rows, # of columns)*
- *A summary of 8-10 variables in the replication dataset. For numeric variables, the summary should include the mean, standard deviation, and count of NAs. For categorical variables, the summary should include each level present in the data and its count, as well as a count of NAs. If multiple datasets are submitted as part of your work, at least one variable should be included from each dataset.*

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.*

The R script can be found here.

The files needed to run the R code are:

(1) "continent.csv"
- Contains the country name and the continent which it belongs to
- Will need to create a "north" dummy in the same way the authors do
  - North = 1 if country is in Europe or North America
  - Move Mexico to the South (Note: Turkey is already in Asia)
  - Move Australia, New Zealand, Israel, Japan, and South Korea to the North

(2) "gdp.csv"
- Contains the country name, year, and GDP in million USD

(3) "population.csv"
- Contains the country name, year, and population in million persons

(4) “unemployment.csv”
    - Contains the country name, year, and unemployed as a percent of labor force
(5) “oecdratification.csv”
    - Contains the country name, exact date a country became and OECD member country, and which countries are members of the 18 countries the original authors have done their analysis on
(6) All files under the trade folder
    - Contains bilateral trade between an OECD member country and any country in the world

The steps in merging all csv files are:

(1) Import all non-trade data
(2) Create a “North” dummy using continents data
(3) Import trade data
(4) Merge trade data with continents data
(5) Merge trade and non-trade data
(6) Create file without missing observations

There are 3 final replication datasets of interest:

(1) finaldata_withflags: contains final data and additional flag columns
(2) finaldata: contains final data without the flag columns
(3) finaldata_noNA: contains the final data, after dropping all rows with NAs

The final replication dataset has all 33 OECD countries from 1950-2011. There are 3 possible datasets for the data analyst:

- Drop countries not part of the original 18 countries that the original author used (use the variable “oced18” to do this)
- Drop countries for the years before they became OECD countries (use the variable “year_ratified” to do this)
- Use all 33 countries from 1950-2011

The data analyst should create the yearly variable for national affluence (real GDP / national population). The “flag columns” are for flags created by OECD for countries with values (GDP, population, etc.) that are estimated (“E”).

(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.

The data dictionary is available here. This dictionary is for the data frame “finaldata_withflags”. The data frame “finaldata” will have exactly the same dictionary but less variables while the data frame “finaldata.noNA” drops all NA observations from “finaldata”.

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.***

Data is from 1965 to 2018 (54 years) from OECD countries (33 remaining OECD countries after removing those in the South). Before dropping missing observations, there are 1,765 observations in total. After dropping missing observations, there are 784 observations left.

Required sample size [to be filled out by the SCORE team]: The primary unit of analysis is the year grouped within country. An estimate of the minimum viable sample size for the data analytic replication is: 29. For comparison, the stage1 required sample size would be: 139 and the stage2 sample size would be: 311.

Notes: The SER method used assumes that the number of observations in groups/clusters is the same in the original and replication, so sample sizes should be changed by altering the number of cluster/grouping variables, rather than the number of observations per cluster/group.

14. Sample size rationale

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: https://osf.io/2cefu/?view_only=2b7c42b78c1f4d28a7d685f6420b2804

15. Stopping rule (provided by SCORE)

The SCORE team recommends that three analyses be performed for this replication study:

- A focal analysis that only uses country-years that were not used in the original analysis.
- A second analysis that uses all available country-years.
- A third analysis that only uses country-years that were used in the original analysis.

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?*

DATA SET CHOSEN

There were three data sets provided to choose from in order to perform the focal analysis. The set that was chosen is called "**finaldata_noNA**" (refer to section 12(d) for a description of this specific data set). There are two reasons this set was chosen:

1. Since it is a panel set to be used in a fixed effects estimation it would be best to use complete observations with accurate and available values for each dimension of the observations (i.e.: value for gdp, pop, unemp, totalimports, totalexports available for all years used).
2. It would save a couple of lines of code on the script for the estimation routines as I would not have to direct STATA in how to handle observations which contained an NA instead of an actual value. Further, the manuscript didn't specify how the author would have handled data with missing observations.

VARIABLE NAME

Naff: This variable corresponds to "National Affluence". In the paper, it is defined on page 1652 as *"a country's gross domestic (GDP) divided by its total population, with GDP expressed in U.S. dollars at prices and purchasing power parity (PPP) from the year 2000."* I originally wanted to use the nomenclature "NA" as used by the paper in equation (1b) on page 1659, but this might lead to programming issues since NA is a common command in many statistical data analysis packages. According to section 12 (d) of the pre-registration, gdp and population in the data file are expressed in millions, thus NAff is defined as gdp/pop. There is no definitional deviation for this variable between the original study and replication study.

IMS: This variable corresponds to "Imports From the South". In the paper, it is defined on page 1654 as *"the annual monetary value of imports…" from "...Africa, Asia, Central and South America, and Oceania." Additional adjustments are made "...by moving Mexico and Turkey to the South, and...moving Australia and New Zealand (from Oceania) and Israel, Japan, and South Korea (from Asia) to the North...To facilitate international comparison, values for imports...are expressed as percentage of GDP for all countries"*. According to section 12 (d) of the pre-registration, under the heading Imports from/Exports to the South, the values are in level units. Since GDP is reported in millions, in order to define IMS as a percentage of GDP, IMS is

defined as totalimport/(gdp*10,000). There is no definitional deviation for this variable between the original study and replication study.

EXS: This variable corresponds to “Exports to the South”. Similarly defined as IMS, also on page 1654 of the paper, it is “the annual monetary value of exports...to Europe and North America...as percentage of GDP for all countries.” As mentioned in the definition for IMS, Mexico and Turkey are categorized as Southern countries, and Australia and New Zealand Israel, Japan, and South Korea are moved to the North category. EXS is defined as totalexport/(gdp*10,000). There is no definitional deviation for this variable between the original study and replication study.

Control variables include unemployment and time dummies in order to control for time fixed effects.

unemp: This variable corresponds to the civilian unemployment rate. It is used in order to “account for cross-national variation’’...resulting from “industrial decline...generates rising unemployment rates as well as increased service sector employment” (pg. 1655). The dataset expresses this variable in percentage form already and thus no change in this variable was required. There is no definitional deviation for this variable between the original study and replication study.

DUMXXtoYY: This variable corresponds to a time dummy which indicates 1 if the year of the observation falls between year XX and year YY and zero otherwise. In the original study, the time dummies span 5 years, which is also repeated in the replication. So, for example, if the observation is for 1997, then for the variable DUM95to99, there would be a 1, since 1997 lies in the span 1995 - 1999. For all the other time dummies, this observation would have a value of 0. In the original study, the data used by the author span 1970 through 2003, and the author accounts for “unmeasured, time-specific effects” by introducing “...dummy variables for each 5-year period in the data set, with the period 1970-1974 as the reference category” (page 1657).

While there is no definitional deviation as I also use 5 year dummies for the replication, I use the extended data set provided to me which spans 1967 - 2018, with the period 1967-1969 as the reference category. Thus, the dummies I use include

DUM75to79, DUM80to84, DUM85to89, DUM90to94, DUM95to99, DUM00to04, DUM05to09, DUM10to14, DUM15to18

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?*

There are no composite measures used for the focal analysis.

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 statement confirms that only 25% of the data have been randomly sampled in developing the analysis plan and code contained in this preregistration.

Following the procedure outlined and discussed in the Kollmeyer paper on pages 1656 - 1658, I dispatch STATA's "fixed effects" routine which takes into account the "heterogeneity bias" which occurs as a result of "...unmeasured, country-specific effects" and "...unmeasured time-specific effects" (pgs. 1656 - 1657).

The STATA routine used for the focal analysis can be found by clicking here: (LINK). The routine imports the data set provided by the "data sourcer" which is an (annual frequency) panel data set containing the GDP measure, population, total imports from the South, and total exports to the South for each country spanning the years 1967 through 2018. The .do file assigns an index/I.D. to the panel data (in this case, the name of the country) and then assigns the time index, as required (in this case, yearly, in accordance with the frequency of the data).

The .do file then defines the variables as required to perform the analysis, as defined in section 17 of this pre-registration. National Affluence, "NAff" is defined as per capita GDP, gdp/pop, Imports from the South, "IMS" is defined as total imports from the Southern countries as a percentage of GDP, totalimport/(gdp*10,000), and Exports to the South, "EXS" is defined as total exports to the southern countries as a percentage of GDP, totalexport/(gdp*10,000).

Following the steps outlined in the original paper, outliers are eliminated from the dataset by "employing the Hadi robust outlier detection algorithm" (page 1658). In STATA, this is accomplished with the *hadimvo* command which seeks statistical outliers in the named variables and generates an additional indicator variable in which outliers are assigned a 1 and are 0 otherwise. The .do file then drops these outliers from the dataset.

A "cleaning" stage is then executed by the .do file in which unnecessary columns of data are dropped from the dataset. The routine would still run fine without this step, but eliminating unnecessary columns was a pre-emptive step taken in order to make debugging and other error detection methods easier.

Then, following the paper, a series of dummy variables is constructed in order to truly control the estimation for two-way fixed effects (one-way corresponds to just the group-wise country control, two-way corresponds to both country and time). The .do file does this by first generating the dummy variable and assigning a name and a value of zero to this variable, and then if the observation satisfies the criterion - namely by occurring during a year within the 5 year range specified by the code - it is assigned a value of 1.

After the time indicators are all generated, the .do file then instructs STATA to take a **25% random sample** of the data. The reason 25% is taken instead of the usual 5% is explained below in the following section.

The .do file then instructs STATA to re-sort the panel data set chronologically, since the random sampling ends up scrambling the panel data resulting in a sample which is out of order chronologically. In order to lag the explanatory variables following the original paper "...because cause-and-effect relationships between shifting macroeconomic conditions and firm-level employment decisions are often subject to time delays", the "L.x" STATA command is used in the estimation commands, where "x" is the variable to be lagged; this command requires the data to be sorted chronologically.

Finally, the .do file performs the fixed effects estimation. Two different routines are performed in order to provide robust estimation results. The paper references the "xtgls" STATA command, but this command does not have fixed effects as an option. Unlike the "xtgls" command, the "xtreg" command in STATA does allow for fixed effects (by adding the "fe" command as an argument). The xtreg command then estimates the model consistent with equation 1b in the text (pg. 1659):

NAff = constant + parameter1*IMS + parameter2*EXS + "CV" + error,

controlling for fixed effects amongst the named group: countries. Contained within the "CV" (control variables) are the time indicators, making this estimation a true two-way fixed effects estimation. Along with the time indicators, the other control variable is the unemployment rate.

The .do file then re-estimates the model using generalized least squares, as in the paper. Since fixed effects is not an option in the "xtgls" command, I force the estimation to control for fixed effects amongst countries and time by also providing dummy variables for countries along with the 5 year time segment dummies. This is accomplished by adding "i.countrynum" which provides indicator variables for each country; thus if the observation is for country "x", then a 1 is assigned to the indicator for country "x" with 0s assigned for all of the other indicators. As expected, this estimation provides the **same results** as the "xtreg" command with the "fe" argument.

A snapshot of the results of both estimation routines are provided here: (LINK for the “xtreg” results) and here: (LINK for the first segment of “xtgls” results and LINK for the second segment).

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?*

The only transformation performed by the script file linked in the previous section is a lagging of the explanatory variables. In STATA, the easiest way to lag a variable is to simply declare “L.x” where “x” is the variable you wish to lag. However, in the context of this focal analysis, this results in a couple of important complications. If one is controlling for country fixed effects and there is only one observation for a specific country, if you lag the explanatory variables (imports from the south “IMS”, for example) then STATA will recognize this as a missing variable and will throw out an error and not complete the estimation. This is precisely what happened with a handful of countries in the original data set. For example, Bulgaria, only had data available from 1999 through 2018. If we are only to take a 5% sample, this would always result in a single observation for this country leading to the previously mentioned error.

The way to get around this is to increase the sampling to 25% - this still poses problems for the estimation routine, as that only provides a couple of observations for some of the countries. Performing an estimation on a couple of observations is not going to result in any meaningfully powerful inference, but it does allow the estimation to run without any errors.

Needless to say, the extended data set provided (1967 - 2018) was absolutely required to get around this. If I were to stick with the dataset which matches what was used in the original paper (1970 - 2003), there would simply not be enough observations in order to take a 5% sample (or even 15%) to have the estimation run without errors.

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 and statistically significant (to the 5% level) association between national affluence "NAff" and imports from the south "IMS" as mentioned in section 4 of this pre-registration.**

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?)*

In order to eliminate outliers from the data set, I followed the original paper by deploying the "Hadi robust outlier detection algorithm available in STATA" (page 1658). This algorithm is executed by declaring the "hadimvo" command in STATA. The command is followed by a list of variables you would like to test for outliers. In this study's dataset, the variables national affluence NAff, imports from the south IMS, exports to the south EXS, and the unemployment rate are tested for any outliers. The original paper doesn't mention any parameters, so I allowed

for the default 5% argument in the algorithm. The hadimvo command then generates a new variable, which I called “bad” (for obvious reasons) in which a “1” would be assigned to an observation which is deemed statistically an outlier (to the 5% level of significance) and a “0” otherwise.

As mentioned in section 19, the observations assigned with a 1 by the Hadi outlier detection algorithm are dropped prior to the estimation.

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?*

Having an unbalanced panel is definitely not something strange in econometrics and in fact, may be more common than having a balanced panel. Thus, missing observations are not really an issue when it comes to using statistical methods for panel data when one controls for fixed effects in the groups which have different numbers of observations available.

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.*

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.