**Template**

**Title: A comment on Speedy Analyst et al. (2022)**

Abel Brodeur (University of Ottawa)

[1][2]

**Abstract**

**Instructions**:

*Summarize in few sentences the original study, focusing on the main results in the original abstract in terms of word claim[[1]](#footnote-1) which you attempt to reproduce or replicate. Provide information, if relevant, on the magnitude and statistical significance of the main results. Then report all your reproduction and replication results.*

*In the event that there are too many robustness tests per claim to report them individually in the Abstract, then report a summary measure such as the fraction of tests that replicates (i.e., statistically significant in the same direction as the original result) for each claim and the average relative size of the tests for each claim (or the average of these measures across the clams if there are also too many claims being reproduced or replicated to report them individually in the Abstract).*

*See below definitions:*

*Computational reproducibility: The ability to duplicate the results of a prior study using the same data and procedures as were used by the original investigator. Reproducibility is done using the same computer code, but can be achieved using a different software package.*

*Robustness replicability: The ability to duplicate the results of a prior study using the same data but different procedures as were used by the original investigator. Robustness replicability can be done using the raw, intermediate or final data sets used by the original authors.*

*Direct replicability: The ability to duplicate the results of a prior study using new data but the same procedures as were used by the original investigator.*

*Conceptual replicability: The ability to duplicate the results of a prior study using new data and different procedures as were used by the original investigator.*

**Example**:

Speedy Analyst et al. (2022) examine the effect of a policy implemented in the fictional country Labas. In their preferred analytical specification, the authors find that the policy (PROSCOL) increased educational attainment of the treated group by 33 percentage points and decreased fertility by 9.8%. Their point estimates are statistically significant at the 5% and 10% levels, respectively. First, we reproduce the paper’s main findings and uncover two minor coding errors which have no effect on the studies’ main results. Second, we test the robustness of the results to (1) adding more years to the sample and (2) changing how standard errors are clustered. We find that adding more years to the sample decreases the size of the point estimate by one-third for education and by one-fourth for fertility. The point estimate for fertility becomes statistically insignificant at the 10% level, while it remains significant at the 5% level for education. Clustering at the region level makes the point estimates for education and fertility to be statistically insignificant at the 10% level.

Example for a direct replication of an experimental study: We conduct a direct replication of the paper by using the same procedures (i.e., method and analysis) and new data. We confirm the sign, magnitude and statistical significance of the point estimates for outcome X.

**1. Introduction**

**Instructions**:

*Briefly describe the main data sources, method, policy or treatment, time period and population for which the estimates apply. Then describe the main scientific claims (descriptive or causal) and robustness checks if those that are relevant for your re-analysis or replication. Quote the original part of the study that has the main scientific claim(s) including page number(s). As suggested in the Guide for Accelerating Computational Reproducibility in the Social Sciences (*[*https://bitss.github.io/ACRE/*](https://bitss.github.io/ACRE/)*), structure your summary of the main findings and methodology as follows: "The paper tested the effect of X on Y for population P, using method M. The main results show an effect of magnitude E (specify units and standard errors)" or "The paper estimated the value of Y (estimated or predicted) for population P under dimensions X using method M. The main results presented an estimate of magnitude E (specify units and standard errors)". This template assumes that the paper’s scientific claims are focused on estimating a causal relationship. In the event that the original study is estimating/predicting a descriptive statistic of a population, or something else, then describe and quote the results accordingly using precise claims from the original study.*

*How to select claims to reproduce/replicate? There are three possibilities; (1) select claims for all "hypotheses tests" in the original study, (2) select claims mentioned in the abstract or (3) select claims for what is considered the main result in the paper as stated by the original author(s). For the last option, provide a quote from the original study confirming that the claim chosen is considered the main result by the original author(s).*

*Next, summarize your reproduction and/or replication. Start by stating how you have obtained the data and codes and if the original author(s) answered your request(s) a/nd questions. Indicate the repository where your programs and data are located. Then proceed with a description of your* [*conceptual reproducibility*](https://i4replication.org/definitions.html) *by describing if you have found coding error(s) and how they affect the main conclusions.*

*For replication, we adopt the definitions* [*here*](https://i4replication.org/definitions.html)*. For robustness replicability, clearly state your robustness checks and how they affect the main point estimates. For direct and conceptual replications, clearly describe the new data. For conceptual replications, also describe the new procedures or they differ from the original study. Different procedures implies a different experimental design and/or analysis for experiments and different methods and/or analysis for observational data.*

*For all three replication types, be precise and summarize your results as follows: “Implementing this robustness increases/decreases the size of the main point estimate for outcome Y by X and the estimate is not anymore statistically significant at the X% level ” or "Implementing this robustness check has no effect on the magnitude or the statistical significance of the main point estimate." Also report the coefficient (or other effect size), the standard error of the coefficient/effect size, the test statistic including df if relevant, and the p-value for all tests.*

**Example**:

Ms. Speedy Analyst, Mr. Speedy Analyst and John Doe (2022), henceforth SSD, investigate the impact of a program called PROSCOL. The setting is the country of Labas. In 2000, its government introduced an antipoverty program in northwest Labas (Ravallion, 2001). The program aims to provide cash transfers to poor families with school-age children. To be eligible to receive the transfer, households must have certain observable characteristics that suggest they are poor.

SSD tested the effect of the policy (PROSCOL) on school enrollment and fertility for low-income families, using a difference-in-differences approach comparing the treated region (Labas) to untreated regions before/after the implementation of the policy. The main data set comes from the Labas Social Survey from 1998 to 2002. SSD’s describe their main results on p.7 as follows “we show that the policy (PROSCOL) increased school enrollment rates for the treated group by 30 percentage points and decreased the number of children born by 0.10 per family (mean of the dependent variable is 3.4). Our point estimates are statistically significant at the 5% level.”[3]

In the present paper, we investigate whether their analytical results are reproducible and replicable and further test their robustness to two specification checks: (1) adding more years to the sample and (2) changing how standard errors are clustered. In their original analysis, SSD rely on data from 1998 to 2002 and cluster their standard errors at the region/year level. In our re-analysis, we extend the time period to 1998 to 2004 and cluster the standard errors at the region level. We are grateful to the original authors for providing these additional years of data, which were (un)available at the time of their analysis.

In terms of reproducibility, we would like to acknowledge that the original study was successfully reproduced by the data editor’s team at the *American Economic Review*. We also successfully reproduced SSD’s main tables (Tables 4 and 5) using their codes, although there were very small discrepancies in the magnitude of the main point estimates for Table 5 due to coding errors. We uncovered two minor coding errors; (1) coding the control variable Age and (2) the gender dummy was included as a continuous variable in one regression.

We then turn to sensitivity analysis. As mentioned above, we test the robustness of the results to (1) adding more years to the sample and (2) changing how standard errors are clustered. We find that adding more years to the sample decreases the size of the main point estimate by one-third for educational attainment and by one-fourth for fertility. The point estimate for fertility becomes statistically insignificant at the 5% level, while it remains significant at the 5% level for education. Clustering at the region level makes the point estimates to be statistically insignificant at the 5% level.

Last, we attempt to replicate the paper using the raw data and new codes. We would like to thank the original authors for making available the raw data; educational attainment, fertility, demographic data and PROSCOL data.

**2. Reproducibility**

**Instructions**:

*Describe in details if you have found coding error(s) and how they affect the main conclusions. In the event that the sign, magnitude or statistical significance is changed for the main estimates or robustness checks, report the new point estimates (e.g., in a table) in this section.*

*Fix the coding error(s) prior to conducting the replication, but make sure to clearly state and disentangle the effect of the coding error(s) vs the change(s) made to the data and codes/procedures in Section 3 - Replication.*

**Example**:

We describe in this section two minor coding errors that we uncovered while reproducing the study. First, we noticed that the coding of the control variable Age was incorrect. Age was defined as the age of the mother in the paper but coded using the variable age of the head of the household in the codes. Second, the gender dummy was included as a continuous variable in one regression. Our codes/programs are available here (e.g., OSF webpage with DOI).[4] The original authors’ updated codes are available here (e.g., OSF webpage with DOI).

We re-run the codes correcting these two errors and reproduce the results for the outcome variable fertility in Table 1. (The specification for educational attainment does not include these control variables.) The structure of the table is the same as in the original study. We find that the point estimates are strikingly similar, with the sign, magnitude and statistical significance being remarkably similar.

**3. Replication**

**Instructions**:

*Clearly state/describe which type of replication you are conducting. See definitions at the beginning of this document. For robustness replicability, present your robustness checks and how they impact the main point estimates one by one so that it is clear how each modification to the specification/analysis impacts the main conclusions. Then you may combine them. Also, clearly state why you conduct each specific robustness check and/or modify the setting/model.*

*Do not confuse general critiques of the original research with replication or robustness checks (Brown and Wood, 2018). For instance, any critique of the design or methodology (e.g., qualitatively discussing the validity of an exclusion restriction for an instrumental variable) should be included in a separate section and clearly labeled as general critiques rather than a replication exercise.*

**Example**:

We now turn our attention to our replication. We test the robustness of the results to a direct replication by adding more years to the sample and a robustness replicability by changing how standard errors are clustered. We add more years to the sample to increase the sample size as the original study was underpowered (e.g., 40% power). We cluster the standard errors at the region level instead of at the region/year level to account for non-independence between years within each region.

The decision to conduct these two robustness checks was taken after reading the paper but prior to observing the codes/programs.[5] We pre-registered our sensitivity analysis here (hyperlink to your pre-analysis plan).

**3.1 Regression model**

For our analysis, we rely on the same specifications and a difference-in-differences analysis comparing the treated region (Labas) to untreated regions before/after the implementation of the policy, restricting the sample to low-income families.[6] The analysis is at the region/year level for the educational attainment outcome and at the family level for the fertility outcome. See the original study for more details and equations. See an example of what to write if the specification, model or method is different in Appendix I.

**3.2 Results educational attainment**

**3.2.1 Extending the time period**

We first investigate whether extending the time period to 2004 has an impact on the sign, magnitude and statistical significance of the difference-in-differences model for educational attainment. For this analysis, an observation is a region/year. The sample is restricted to low-income families and collapsed at the region/year level. There are 20 regions and seven years. The dependent variable, educational attainment, is the fraction of low-income families with all school-age children attending school in a region in a year. Robust standard errors are in parentheses, clustered by region/year. Our findings are reported in Table 2. Column 1 reports the preferred estimates from the original study.[7]

The preferred specification, as determined by the original authors, is presented in column 2. Column 3 adds control variables as in the original study. We find that the policy (PROSCOL) increased school enrollment rates for the treated group by 21 percentage points (in comparison to 30 percentage points in the original study). The point estimate remains statistically significant at the 5% level.

**3.2.2 Clustering**

We then investigate whether changing the clustering technique affects the main point estimates for educational attainment. Robust standard errors are in parentheses, clustered by region are reported in Table 3. Recall that the original study clustered by region/year. Column 1 reports the preferred estimates from the original study. We find that the point estimate for the preferred specification (column 2) is not anymore statistically significant at the 5% level. The standard error is now much larger (0.235) than in the original study (0.151).[8]

**3.3 Fertility**

**3.3.1 Extending the time period**

We turn to the second outcome variable, fertility, in this subsection. We first extend the time period to 2004. For this analysis, an observation is a family. The sample is restricted to low-income families. The dependent variable is the number of children and the mean of the dependent variable is 3.3 (3.4 in the original study). Robust standard errors are in parentheses, clustered by region/year. Our findings are reported in Table 4. Column 1 reports the preferred estimates from the original study.

We find that adding more years to the sample decreases the size of the main point estimate by one-fourth; the point estimate in column 2 (preferred specification) is now -0.075, while it was -0.098 in the original study. The point estimate becomes statistically insignificant at the 5% level (standard error 0.058).

**3.3.2 Clustering**

We now turn to changing the clustering technique for fertility. Robust standard errors are in parentheses, clustered by region are reported in Table 5. We find that the point estimate for the preferred specification (column 2) is not anymore statistically significant at the 5% level. The standard error is now much larger (0.075) than in the original study (0.055).

**3.4 Extending time period and clustering**

We now explore the effect of changing the time period and clustering simultaneously on the two main dependent variables. The estimates are presented in Tables 6 and 7. Overall, we confirm our previous conclusions.

**4. Conclusion**

*State the most important results of your work and what you have learned. You may also describe other empirical exercises that could be conducted by other replicators.*

**References**

Analyst, S., Analyst S. and Doe J. (2022). The impact of PROSCOL on educational attainment and fertility. The Journal, 1(1), 1-10.

Brown, A. N. and Wood, B. D. K. (2018). Which tests not witch hunts: a diagnostic approach for conducting replication research. Economics, 12(1), 20180053.

Ravallion, M. (2001). The mystery of the vanishing benefits: An introduction to impact evaluation. The World Bank Economic Review, 15(1), 115-140.

**Figures**

Insert figures here.

**Tables**

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1 – Coding error: Fertility** | | | |
| **OLS** |  | **Number of**  **Children**  **Original Study**  **(1)** | **Number of**  **Children**  **(2)** |
| Treated (Labas) |  | 0.279\*\*\* | 0.278\*\*\* |
|  |  | (0.079) | (0.079) |
| After Policy |  | 0.047 | 0.047 |
|  |  | (0.045) | (0.045) |
| Treated\*After |  | -0.100\*\* | -0.101\*\* |
|  |  | (0.042) | (0.042) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls |  | Yes | Yes |
| N |  | 12311 | 12311 |
| R² |  | 0.100 | 0.100 |
| Note: Authors’ calculations using data from 1998-2002. Mention data sources here. An observation is a family. The sample is restricted to low-income families. There are 20 regions and five years. The dependent variable is the number of children. Robust standard errors are in parentheses, clustered by region/year.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 2 – Impact of the treatment on educational attainment: Changing time period** | | | |
| **OLS** | **Educational**  **Attainment**  **Original Study**  **(1)** | **Educational**  **Attainment**  **(2)** | **Educational**  **Attainment**  **(3)** |
| Treated (Labas) | -0.147\*\*\* | -0.168\*\*\* | -0.208\*\*\* |
|  | (0.020) | (0.024) | (0.041) |
| After Policy | 0.041 | 0.043 | 0.050 |
|  | (0.030) | (0.029) | (0.047) |
| Treated\*After | 0.333\*\* | 0.210\*\* | 0.200\*\* |
|  | (0.151) | (0.095) | (0.097) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 100 | 140 | 140 |
| R² | 0.047 | 0.048 | 0.051 |
| Note: Authors’ calculations using data from 1998-2004. Mention data sources here. An observation is a region/year. The sample is restricted to low-income families and collapsed at the region/year level. There are 20 regions and seven years. The dependent variable, educational attainment, is the fraction of low-income families with all school-age children attending school in a region in a year. Robust standard errors are in parentheses, clustered by region/year.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3 – Impact of the treatment on educational attainment: Changing clustering** | | | |
| **OLS** | **Educational**  **Attainment**  **Original Study**  **(1)** | **Educational**  **Attainment**  **(2)** | **Educational**  **Attainment**  **(3)** |
| Treated (Labas) | -0.147\*\*\* | -0.147\* | -0.188\* |
|  | (0.020) | (0.085) | (0.103) |
| After Policy | 0.041 | 0.041 | 0.047 |
|  | (0.030) | (0.045) | (0.052) |
| Treated\*After | 0.333\*\* | 0.333 | 0.300 |
|  | (0.151) | (0.235) | (0.216) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 100 | 100 | 100 |
| R² | 0.047 | 0.047 | 0.050 |
| Note: Authors’ calculations using data from 1998-2002. Mention data sources here. An observation is a region/year. The sample is restricted to low-income families and collapsed at the region/year level. There are 20 regions and five years. The dependent variable, educational attainment, is the fraction of low-income families with all school-age children attending school in a region in a year. Robust standard errors are in parentheses, clustered by region.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 4 – Impact of the treatment on fertility: Changing time period** | | | |
| **OLS** | **Number of**  **Children**  **Original Study**  **(1)** | **Number of**  **Children**  **(2)** | **Number of**  **Children**  **(3)** |
| Treated (Labas) | 0.375\*\* | 0.275\*\* | 0.249\*\* |
|  | (0.157) | (0.127) | (0.119) |
| After Policy | 0.041 | 0.031 | 0.037 |
|  | (0.030) | (0.025) | (0.035) |
| Treated\*After | -0.098\* | -0.075 | -0.075 |
|  | (0.055) | (0.058) | (0.057) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 15345 | 15345 | 15300 |
| R² | 0.098 | 0.098 | 0.100 |
| Note: Authors’ calculations using data from 1998-2004. Mention data sources here. An observation is a family. The sample is restricted to low-income families. There are 20 regions and seven years. The dependent variable is the number of children. Robust standard errors are in parentheses, clustered by region/year.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 5 – Impact of the treatment on fertility: Changing clustering** | | | |
| **OLS** | **Number of**  **Children**  **Original Study**  **(1)** | **Number of**  **Children**  **(2)** | **Number of**  **Children**  **(3)** |
| Treated (Labas) | 0.375\*\* | 0.375 | 0.279 |
|  | (0.157) | (0.357) | (0.289) |
| After Policy | 0.041 | 0.041 | 0.047 |
|  | (0.030) | (0.037) | (0.045) |
| Treated\*After | -0.098\* | -0.098 | -0.100 |
|  | (0.055) | (0.075) | (0.073) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 12345 | 12345 | 12311 |
| R² | 0.098 | 0.098 | 0.100 |
| Note: Authors’ calculations using data from 1998-2002. Mention data sources here. An observation is a family. The sample is restricted to low-income families. There are 20 regions and five years. The dependent variable is the number of children. Robust standard errors are in parentheses, clustered by region.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 6 – Impact of the treatment on educational attainment: Changing time period and clustering** | | | |
| **OLS** | **Educational**  **Attainment**  **Original Study**  **(1)** | **Educational**  **Attainment**  **(2)** | **Educational**  **Attainment**  **(3)** |
| Treated (Labas) | -0.147\*\*\* | -0.168\* | -0.208\* |
|  | (0.020) | (0.085) | (0.103) |
| After Policy | 0.041 | 0.043 | 0.050 |
|  | (0.030) | (0.045) | (0.052) |
| Treated\*After | 0.333\*\* | 0.210 | 0.200 |
|  | (0.151) | (0.235) | (0.216) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 100 | 140 | 140 |
| R² | 0.047 | 0.048 | 0.051 |
| Note: Authors’ calculations using data from 1998-2004. Mention data sources here. An observation is a region/year. The sample is restricted to low-income families and collapsed at the region/year level. There are 20 regions and seven years. The dependent variable, educational attainment, is the fraction of low-income families with all school-age children attending school in a region in a year. Robust standard errors are in parentheses, clustered by region.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 7 – Impact of the treatment on fertility: Changing time period and clustering** | | | |
| **OLS** | **Number of**  **Children**  **Original Study**  **(1)** | **Number of**  **Children**  **(2)** | **Number of**  **Children**  **(3)** |
| Treated (Labas) | 0.375\*\* | 0.275 | 0.249 |
|  | (0.157) | (0.357) | (0.289) |
| After Policy | 0.041 | 0.031 | 0.037 |
|  | (0.030) | (0.037) | (0.045) |
| Treated\*After | -0.098\* | -0.075 | -0.075 |
|  | (0.055) | (0.075) | (0.073) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 15345 | 15345 | 15300 |
| R² | 0.098 | 0.098 | 0.100 |
| Note: Authors’ calculations using data from 1998-2004. Mention data sources here. An observation is a family. The sample is restricted to low-income families. There are 20 regions and seven years. The dependent variable is the number of children. Robust standard errors are in parentheses, clustered by region/year.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

**Appendix**

**Appendix I – Example of what to write if the replicator(s)’ specification, model or method differs from the original study for the sensitivity analysis**.

We test the claim that PROSCOL impacted enrollment rates and the number of children born for treated regions in comparison to control regions using a triple differences model comparing low-income and high-income families in treated and control regions before/after the implementation of the policy. This setting contrasts with the original study which relies on a difference-in-differences comparing treated regions to untreated before/after the implementation of the policy for low-income families. Our equation is:

*Y*irt = *α* + *β*1 *Treated*rt + *β*2 *Low Income*irt + *β*3 *After*t + *β*4 *Treated*t × *Low Income*irt + *β*5*Treated*rt× *After*t + *β*6 *Low Income*irt× *After*t + *β*7*Treated*rt× *Low Income*irt× *After*t + ζ*X*irt + εit

where the dependent variable is the number of children born for individual *i* in region *r* in time *t*. The time period is 1998-2004 (1998-2002 in the original study) and an observation is a family (head of the family for the control variable). Treated region is an indicator which is set to 1 if the respondent lives in the treated region (northwest region of Labas), Low Income indicates whether the family is low-income, and After is a variable that takes the value 1 if the year is after 1999. The interaction of Low Income and Treated Region determines the treated group, while the additional interaction with After shows the effect of the treatment. The coefficient of interest is *β*7. *X* represents a vector of family socioeconomic variables (head of the family). We use robust standard errors and cluster the error term at the region level (region/year level in the original study).

**Appendix II – Reproduction of the main tables (Tables 4 and 5) from the original study.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 4 – Original Study: Educational Attainment** | | | |
| **OLS** | **Educational**  **Attainment**  **(1)** | **Educational**  **Attainment**  **(2)** | **Educational**  **Attainment**  **(3)** |
| Treated (Labas) | -0.129\*\*\* | -0.147\*\*\* | -0.188\*\*\* |
|  | (0.021) | (0.020) | (0.039) |
| After Policy | 0.020 | 0.041 | 0.047 |
|  | (0.017) | (0.030) | (0.045) |
| Treated\*After | — | 0.333\*\* | 0.300\*\* |
|  |  | (0.151) | (0.129) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 100 | 100 | 100 |
| R² | 0.047 | 0.047 | 0.050 |
| Note: Authors’ calculations using data from 1998-2002. Mention data sources here. An observation is a region/year. The sample is restricted to low-income families and collapsed at the region/year level. There are 20 regions and five years. The dependent variable, educational attainment, is the fraction of low-income families with all school-age children attending school in a region in a year. Robust standard errors are in parentheses, clustered by region/year.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 5 – Original Study: Fertility** | | | |
| **OLS** | **Number of**  **Children**  **(1)** | **Number of**  **Children**  **(2)** | **Number of**  **Children**  **(3)** |
| Treated (Labas) | 0.367\*\* | 0.375\*\* | 0.279\*\*\* |
|  | (0.151) | (0.157) | (0.079) |
| After Policy | 0.020 | 0.041 | 0.047 |
|  | (0.017) | (0.030) | (0.045) |
| Treated\*After | — | -0.098\* | -0.100\*\* |
|  |  | (0.055) | (0.042) |
|  |  |  |  |
| **Control Variables** |  |  |  |
| Demographic Controls | No | No | Yes |
| N | 12345 | 12345 | 12311 |
| R² | 0.097 | 0.098 | 0.100 |
| Note: Authors’ calculations using data from 1998-2002. Mention data sources here. An observation is a family. The sample is restricted to low-income families. There are 20 regions and five years. The dependent variable is the number of children. Robust standard errors are in parentheses, clustered by region/year.  Significant at the \*\*\*[1%] \*\*[5%] \*[10%] level. | | | |
|

[1] For each author: List your affiliation(s) and contact information. Indicate who is the corresponding author if multiple authors.

[2] For each author: Acknowledge any financial support or conflict of interest. Describe your relationship with the original author(s) if there is a conflict of interest. Examples of conflict of interest include, but are not limited to, being a colleague, collaborator, current or former student, former thesis supervisor or family member. See I4R’s conflict of interest policy here:<https://i4replication.org/conflict.html>.

[3] Report the statistical significance used by the original authors.

[4] Make sure to cite and clearly reference your data and codes.

[5] You should be honest about whether your sensitivity analysis was conceived before or after looking at the programs of the original author(s) and state whether your replication or sensitivity analysis was pre-registered.

[6] Make sure to mention the main statistical or econometric method used to examine each claim and whether the method that you use is the same as in the original study. Also, you should state and rely on the original authors’ preferred specification (or yours, if the authors are not clear).

[7] It is useful to offer a direct comparison to the original authors‘ estimates, for instance, in the first column of your tables. Another option is to reproduce the entire table. See the Appendix for a reproduction of the original authors’ point estimates (Table 4 in their paper).

[8] Be careful to not label differences between your estimates and the original author(s) estimates as mistakes or errors. Rather, help the reader better understand why you conducted a specific robustness check or modified the setting/model.

1. By main results, please follow the word claim from the Social Science Replication Platform: “Claims can have different structures, here we propose two high-level categories based on the most common structures:

   Causal claim: a claim is causal if it can be summarize using causal language. This language can be characterize by the following structure: “The paper estimates the effect of a variable X on outcome Y for population P, using method M.” For example: “This paper investigates the impact of bicycle provision (X) on secondary school enrollment (Y) among young women in Bihar/India (P), using a Difference in Difference approach (M).”

   Descriptive/predictive claim: a claim is descriptive or predictive if it can be summarize using descriptive or predictive language. language can be characterize by the following structure: “The paper estimates the value of a variable Y (estimated or predicted) for population P under dimensions X (optional) using method M.” For example, “Drawing on a unique Swiss data set (P) and exploiting systematic anomalies in countries’ portfolio investment positions (M), I find that around 8% of the global financial wealth of households is held in tax havens (Y).” [↑](#footnote-ref-1)