

# Problem Set #1. Student Report on Reproduction Attempt

Econ 270D

4/24/2020

**Student Name:** 4

**Paper to Reproduce:** Railroads of the Raj: Estimating the Impact of Transportation Infrastructure

Link: <http://dx.doi.org/10.1257/aer.20101199>

Number of entries in ACRE (including this one): 2

## Paper Characteristics

Table 1: Number of claims recorded in ACRE for this paper

	Whole paper	Reproduction Attempt
mean	4	2.5
sd	0	2.1

## Type of paper

*Student:* Causal

*ACRE:* Causal (with 50%)

## Reproducibility

**Master Script?:** No

\*\*Reproducibility Improvements Made: \*\*Master script to execute all steps, Standard/clear file organization

## Claim-level Report

Claims are major hypotheses described in a paper that are supported by estimates presented in one or more display items. This section lists the claims the student pre-specified as those to be reproduced, details on how the estimates related to those claims are presented in the paper, and the reproducibility of the display items containing those estimates.

**Summary of Claim 1:** “The paper tested the effect of reduced transportation cost on trade costs for Indian districts that received railway access in 1870-1930 using a non-linear least squares procedure. The results found an elasticity between the two costs of .169.”

Table 2: Main Estimates for Claim 1

	Preferred Specification	Alternative Specification 1
Estimate	.169	.088
SE	-	.028
Units	% change (estimate is an elasticity)	% change (estimate is an elasticity)
Display item	Table 2	Table 2
Page	917	917
Location <sup>a</sup>	[2,2]	[1,1]

<sup>a</sup> For tables [Row,Column]; for inline results approximate paragraph number within page)

Table 3: Display Item<sup>a</sup> Reproducibility Scores for Claim 1

Display Item	Initial Reproducibility Score	Final Reproducibility Score
Table 2	10	10

<sup>a</sup> Display items may contain multiple estimates and is the level at which the reproducibility score is assigned.

## Display Item-level Report

Display items contain the estimates used to support claims and are the unit at which reproducibility is assessed. This section describes the reproducer’s assessment of reproducibility of all display items assessed, including those that do not contain any pre-specified claims.

### Contain Pre-Specified Claims:

Table 4: Main Estimates for Claim 1

	.
Display Item	Table 2
Assessed in entirety?	In its entirety
Complete or incomplete workflow?	Complete
Necessary files to reproduce?	Yes, including the raw data sets.
Able to execute steps in diagram without errors?	No
Output matches publication?	Yes, it is exactly the same.
Proprietary data?	No
Reproducibility score:	Level 10: All materials are available, and the display item can be faithfully reproduced from raw data (pass CRR).

**Contain Claims That Were Not Pre-Specified:**

Table 5: Main Estimates for Claim 1				
	X1	X2	X3	X4
Display Item	Table 1	Table 3	Table 4	Table 5
Assessed in entirety?	In its entirety	In its entirety	In its entirety	In its entirety
Complete or incomplete workflow?	Complete	Complete	Complete	Complete
Necessary files to reproduce?	Yes, including the analytic datasets, but not the raw datasets.	Yes, including the analytic datasets, but not the raw datasets.	Yes, including the analytic datasets, but not the raw datasets.	Yes, including the analytic datasets, but not the raw datasets.
Able to execute steps in diagram without errors?	Yes	No	No	No
Output matches publication?	Yes, it is exactly the same.	Yes, it is exactly the same.	Yes, it is exactly the same.	Yes, it is exactly the same.
Proprietary data?	No	No	No	No
Reproducibility score:	Level 5: Analytic data and analysis code are available, complete and they faithfully produce the display item (pass CRA).	Level 5: Analytic data and analysis code are available, complete and they faithfully produce the display item (pass CRA).	Level 5: Analytic data and analysis code are available, complete and they faithfully produce the display item (pass CRA).	Level 5: Analytic data and analysis code are available, complete and they faithfully produce the display item (pass CRA).

Suggested next steps Next steps: - add information from survey two to df, df\_tidy\_estim, df\_tidy\_claims - bring code to generate df\_tidy\_items from 02\_report\_on\_pilot.Rmd. - explain the data rel between estim, disp items, and claims. Be clear to label DIs that do not have a pre-specified claim or estimate.

For determining which display items are generally associated with the same claims: <https://www.cc.gatech.edu/~vempala/papers/dfkvv.pdf>