# Supplemental Appendix

for

**Comment on "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil"**

*by* JÖRG ANKEL-PETERS, GUNTHER BENSCH, AND COLIN VANCE

## Appendix A. Details on stepwise replication approach

**Step 1:** These results reproduce those of LMB using the replication data that was made available on the *AEJ:AE* website at the time of publication, henceforth LMB (2013*b*). In order to have the same number of observations across all steps, we remove one county from LMB's original dataset, as data for this county is missing to construct the suitability index used in part of the revised specifications. The estimations are therefore based on 8726 instead of 8730 observations at the county-decade level.

**Step 2:** The purpose of this step is to reproduce the output of LMB when re-running their entire analysis using the replication package that the original authors kindly shared (Lipscomb, Mobarak, and Szerman 2022, referred to as LMS in the following). When sharing the simulation code, the original authors noted that they had added a seed to the original code to make the results of the grid simulation reproducible (see also Lipscomb, Mobarak, and Szerman 2021). Additional fine-tuning of the IV simulation model and its inputs appears to have occurred as well. Importantly, we have no reason to believe that any of the changes that may have been made by the original authors have altered the logic of the IV modelling, but rather that they represent improvements in the modelling of the IV in the spirit of LMB. For example, the probit regressions that provide input to the modelling changed slightly. Unlike in the LMB data, the probit regressions in the LMS data do not include those grid points that could not be linked to municipality data, reducing the number of observations from 33,342 to 32,608.

In this step, we find that IV simulation model outputs vary in a consequential manner over multiple runs with different seeds. We therefore run the analysis 100 times in this step and in the following steps, using a set of 100 different seeds to allow for reproducibility. The simulations are run in the programming environment MATLAB® already used by the original authors.

In the results table of the main paper, we present estimates for what we refer to as 'median seed run' as a representative single result from among the 100 seed runs per step. To make sure that the same median seed run is selected for specifications with identical first stages, we select the median seed run across outcomes and steps if they share the same first stage.

**Step 3:** The Amazon is a sizable region in Brazil with low temporal variation in electrification, given the high cost of infrastructure development. This led LMB to adopt a special treatment of the Amazon in the two stages of their analysis of electricity placement in Brazil. First, they use the Amazon as an indicator of the presence of dense forests in the construction of the IV, in order to proxy the high cost of building electricity infrastructure in areas with dense forests. This definition corresponds to the Northern macro-region as defined by the Brazilian statistical agency, IBGE. However, in constructing the control variables of their main specification, they use a more expansive definition of the Amazon including transitional areas, which we call "Extended Amazon". Moreover, both definitions differ from a more accurate, vegetation-based delineation of the Amazon biome as shown in Table A1. We therefore adopt the Amazon biome definition in our analysis, for which we use GIS data shared by the original authors as part of LMS.

We furthermore apply LMB's two old Amazon definitions consistently in two separate robustness checks for Step 3. Results are presented in column (1) and (2) of Table B1 in Appendix B.

**Step 4:** In this step we reconstruct the control and outcome variables and identify a few small but consequential errors and inconsistencies. The raw data is retrieved from the webpage of the *Instituto de Pesquisa Econômica Aplicada*, IPEA (2023). The outcome *in-migration* cannot be retrieved from IPEA, which is why we use the variable included in LMB (2013*b*), recognizing that it is unclear to us where this data comes from. The two main issues relate to variable adjustments required to account for changes in variable definition over time and data aggregation from the municipality to the county level, both of which are explained in detail in the following.

*Variable adjustment*

Some outcomes changed definition in the 1990s. These are the *HDI* and its three sub-components *education*, *longevity*, and *income* as well as the *poverty ratio*, *illiteracy rate*, the *share of people with education less than four years*, and the *years of schooling*. For these outcomes $y$, the

**Table A1: Amazon definitions**

| | Amazon North | Extended Amazon | Amazon Biome | Legal Amazon |
|---|---|---|---|---|
| Map |  |  |  |  |
| Area | 3.9 million km² (46% of the country) | 5.5 million km² (65% of the country) | 5.0 million km² (58% of the country) | 5.1 million km² (60% of the country) |
| Population (2000) | 13 million (8% of total population) | 25 million (15% of total population) | 19 million (12% of total population) | 21 million (12% of total population) |
| Application | LMB in IV simulation model | LMB in 2SLS estimation | Alternative definition used by Lipscomb, Mobarak, and Szerman (2021) and in this comment | Definition not applied beyond sensitivity analyses and not presented as part of this Comment |
| Basis for definition | One of the five so-called macro-regions in Brazil as defined by the *Brazilian Institute of Geography and Statistics* (IBGE), corresponding to seven northern states | Two macro-regions, North and Central-West, that together correspond to ten states including the federal district that contains the capital city, Brasília | Amazon biome as defined by the *Brazilian Institute for the Environment and Natural Resources* (IBAMA) | Amazon as politically defined by Brazilian federal law in 1953, corresponding to nine northern states |
| Advantages and disadvantages | Fails to include jungle-like areas in the Pantanal and parts of the Amazon | Includes extensive parts of non-jungle Cerrado, e.g. Goiás and Brasília | Accounts best for the particularities of rainforest vegetation, at least at one point in time, so without accounting for deforestation | Reflects the region that receives special political treatment to promote protection and development policies |

*Note: Maps created using shapefiles from Django GIS Brasil (2013).*

dataset includes for the decade 1990, $d = 9$, variables with both the old and new definition, i.e. $y_{d=9}^{old}$ and $y_{d=9}^{new}$.

As noted in online appendix 3 of LMB, the data overlap in that decade can be used to adjust data for the year 2000 ($d = 10$) by multiplying it by the following ratio as an adjustment factor:

$$y\_ratio_m = \frac{y_{m,d=9}^{old}}{y_{m,d=9}^{new}}.$$

Here, $m$ refers to the unit of observation of the raw data, which is the municipality. This data is later aggregated at the level of a county, $c$.

LMB made this adjustment for decade 10 to all outcomes listed above except for *HDI*. Furthermore, they used the new-definition variable of the *HDI* in decade 9 as well. In the revision as part of this replication (ABV), we therefore adjust the *HDI* for decade 9 and 10 at municipality level in the same fashion as all other variables with changing definition in the 1990s. This is summarized in Table A2.

For one outcome variable – *income* as HDI sub-component – the adjustment caused the values from 2000 to exceed 1.00, that is, the maximum possible value as defined by UNDP. We therefore top-code these cases to 1.00. This happens with less than five percent of counties. Since the revisions of the three sub-components of the *HDI* need to be reflected in the *HDI* variable, we recalculate *HDI* as (*education* + *longevity* + *income)*/3 at municipality level, that is as the arithmetic mean of its three sub-components, as it was also calculated by IPEA.

**Table A2: HDI variable adjustment**

| | | Variable definition applied | |
|---|---|---|---|
| **Variable** | | LMB | ABV |
| HDI in decade 9 and 10 | $hdi_{m,d=9}$ | $hdi^{new}_{m,d=9}$ | $hdi^{old}_{m,d=9}$ |
| | $hdi_{m,d=10}$ | $hdi^{new}_{m,d=10}$ | $hdi^{new}_{m,d=10} \times hdi\_ratio_m$ |

*Data aggregation*

We also correct a coding error in how LMB aggregated data from the municipality level to the level of analysis, the "Minimum Comparable Areas" known in Portuguese as "Áreas mínimas comparáveis" (AMCs), referred to as counties by LMB. This coding error affects all the eight outcome variables listed above for which the definition changed in the 1990s. Since the number of municipalities increased from about 4500 to about 5500 in that decade, no data from before municipality creation was available for about 1000 municipalities, mostly small municipalities with less than 20,000 inhabitants (Brandt 2010). Hence, the adjustment factor, $y\_ratio$, cannot be derived for these municipalities and their year-2000 variables were set to missing by LMB. LMB then aggregated this data from the municipality to the AMC level by taking population-weighted averages for each outcome. The coding error occurred in that the – partly missing – outcome data entered the numerator, while the population of the entire AMC entered the denominator, instead of only the population of municipalities for which outcome data is non-missing (see the formula in Table A4). This is equivalent to setting the year-2000 data for the eight variables to a real zero for those municipalities with missing adjustment factors.

We illustrate the problem and our approach to dealing with the problem using the data from Table A3. The table presents data on two AMCs from the states of Maranhao and Rio Grande do Sul, where some municipalities were created only in the 1990s (see column 4), by splitting

up from other municipalities in the same AMC. In the AMC with the ID 2232, for example, Novo Machado was split from Tucunduva and Porto Mauá from both Tucunduva and Tuparendi (Estado do Rio Grande do Sul 2018). Since no data for the year 1990 and before is available for the newly created municipalities, the adjustment factors to account for the new post-1990 definitions cannot be calculated (see column 6). In turn, the data at municipality level is missing for the year 2000 for these newly created municipalities (see column 7). For illustration, we use one of the variables that changed definition: *education* as an *HDI* sub-component. Population-weighted AMC level values of this variable should be derived based on the values in column 5 (population) and column 7 (municipality-level outcome data) of municipalities created before 1990 only, shown in capital letters in column (2). However, LMB used also the population data in column 5 of the newly created municipalities. Using the aggregation formula applied by LMB and reproduced in Table A4, LMB calculated the following for the AMC with the ID 2232 (see column 8):

$$\frac{(.803 * 9542) + (.763 * 6305)}{9542 + 2802 + 4718 + 6305} = .534.$$

This, effectively, treats missing values as zeros, as is made explicit through the added second and third summand in the numerator of the following equation:

$$\frac{(.803 * 9542) + (0 * 2802) + (0 * 4718) + (.763 * 6305)}{9542 + 2802 + 4718 + 6305} = .534.$$

Applying the correct formula presented in the "ABV" column in Table A4, we instead arrive at the values shown in column (9):

$$\frac{(.803 * 9542) + (.763 * 6305)}{9542 + 6305} = .787.$$

These values in column (9) are the ones we use in Step 4 of the main paper.

**Table A3: Examples of municipalities with coding errors in data aggregation in LMB**

| | Municipality information | | | | HDI | HDI education (2000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | education | raw | | | imputation-based | |
| AMC60 ID (LMB) | name | code | creation | population (2000) | adjustment factor | data muni. level | LMB AMC level | ABV AMC level | alternative muni. level | alternative AMC level |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| 152 | PIO XII | 2108702 | 1959 | 28413 | .776 | .459 | .332 | .459 | .459 | .457 |
| | Satubinha | 2111722 | 1997 | 10815 | n/a | n/a | | | *.453* | |
| 2232 | TUPARENDI | 4322301 | 1959 | 9542 | .875 | .803 | .534 | .787 | .803 | .780 |
| | Porto Mauá | 4315057 | 1992 | 2802 | n/a | n/a | | | *.774* | |
| | Novo Machado | 4313425 | 1992 | 4718 | n/a | n/a | | | *.759* | |
| | TUCUNDUVA | 4322103 | 1959 | 6305 | .835 | .763 | | | .763 | |

*Note: Municipalities created before 1990 are listed in column (2) with capital letters. n/a refers to not available as the municipality was newly created in the last decade. Values in column (10) based on imputed adjustment factors in italic.*

**Table A4: Aggregation of municipality data**

| Variable | | Variable definition applied | |
|---|---|---|---|
| | | LMB | ABV |
| Aggregated municipality data | $y_{c,d}$ | $\sum_m y_{m=o,d} \times \frac{pop_{m=o,d}}{\sum_m pop_{m=n}}$ | $\sum_m y_{m=o,d} \times \frac{pop_{m=o,d}}{\sum_m pop_{m=o}}$ |
| Variables with changing definition, for which $y\_ratio$ can not be determined | $y_{m=u,d=10}$ | $y^{new}_{m=u,d=10}$ | $y^{new}_{m=u,d=10} \times \overline{y\_ratio}_m$ , with $\overline{y\_ratio}_m = \sum_{-m} y\_ratio_{-m} \times \frac{pop_{-m=o,d}}{\sum_m pop_{-m=o}}$ |

*Note: $m = u$ refers to municipalities with unknown adjustment factor; $m = o$ to municipalities, for which data is available for the respective outcome $y$; $m = n$ represents municipalities with available population data and $-m$ other municipalities in the same county as municipality $m$.*

In the main paper, we mention another approach in Step 4 to deal with the missing values – imputation. We only show results for this alternative approach in the appendix Table B1, also because the first approach discussed above comes closer to what was applied by LMB. Table A4 also shows AMC-level data calculated according to an alternative approach to aggregate the municipality data. The idea behind this approach is to avoid discarding year-2000 information from newly created municipalities by imputing the data instead of setting it to zero. According to this approach, first, the adjustment factors missing for newly created municipalities (see column 6) are imputed by taking the population-weighted mean of the existing adjustment factors of the municipalities created before 1990 from the same AMC (see again column 6). The resulting imputed adjustment factors (not shown in the table) are then used to impute the municipality-level outcome values in column (10), which, in turn, are used to determine the population-weighted mean of the outcome variable (column 11). As can be gleaned from comparing column (9) and column (11), this alternative, imputation-based approach delivers outcome data at the AMC level that differs only marginally from the data calculated according to the first approach described above. The correlations between year-2000 data defined according to the two approaches is always higher than 0.998. Accordingly, it does not come as a surprise that the choice between the two adjustment approaches hardly changes the main estimation results (see column (5) of Table B1).

Both approaches require a similarly strong assumption to deal with the issue of missing data in the year 2000 that occurs because of the newly created municipalities: The first approach assumes that the year-2000 data of municipalities created before 1990 best approximates the data for the entire AMC. The underlying assumption of the alternative approach is that the year-1990 data from municipalities created before 1990 serves to approximate the missing adjustment factors for the newly created municipalities. Essentially, both approaches assume socio-economic similarity between newly created municipalities and the older municipalities that they split from.

Table A5 juxtaposes the county-level data for the outcome *HDI* before and after implementing the two data corrections in terms of variable adjustments and data aggregation. It becomes clear that only data from the third and fourth decade are affected. Nevertheless, this is consequential for *HDI*, as further demonstrated when performing the subsequent Step 5 without having performed Step 4. Related results are presented in column (3) of Table B1 in Appendix B. Additionally, column (4) of Table B1 in Appendix B shows results for the subsequent Step 5 when only the correction of the variable adjustment but not of the data aggregation is made.

**Table A5: County-level means of HDI, before and after data corrections**

| | **mean (sd) of $hdi_{c,d}$** | |
|---|---|---|
| **Decade** | LMB | ABV |
| 1970 | 0.365 | 0.365 |
| | (0.106) | (0.106) |
| 1980 | 0.528 | 0.528 |
| | (0.148) | (0.148) |
| 1990 | 0.626 | 0.567 |
| | (0.098) | (0.145) |
| 2000 | 0.709 | 0.639 |
| | (0.081) | (0.137) |

**Step 5:** LMB already tested the robustness of the estimates to the inclusion of various time-varying controls as alternatives (see table 10 of LMB). While they eventually opted for the Amazon interacted with decade fixed effects, here we argue – in line with the motivation of the earlier corrigendum, Lipscomb, Mobarak, and Szerman (2021) – that one of the alternative controls represents a more comprehensive and flexible way to control for time-varying trends in the geographic factors that affect electricity network expansion: a quartic polynomial in the "suitability index" produced in the probit regressions that proceed the IV construction, as well interacted with decade fixed effects. The "suitability index" not only accounts for Amazon locations, but additionally incorporates information on water flow, river gradient, and land slope, each weighted by their importance in siting decisions for new hydropower plants.

Note also part of the changes outlined above were already incorporated in Bensch, Peters, and Vance (2021) and Lipscomb, Mobarak, and Szerman (2021) as interim replication efforts by our team and the original authors on LMB.

**Additional references**

Brandt, C. T. (2010). A criação de municípios após a Constituição de 1988. *Revista de Informação Legislativa*, *47*(187), 59-75.

Django GIS Brasil (2013). untitled. https://github.com/perone/django-gis-brasil/tree/master/gisbrasil/data/brasil (accessed April 10, 2024).

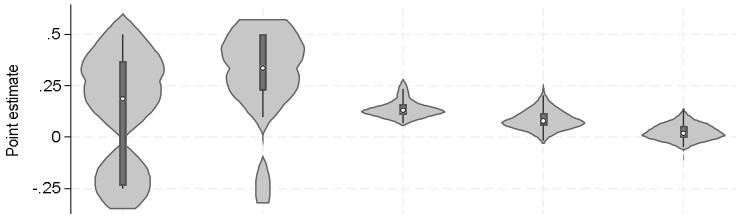
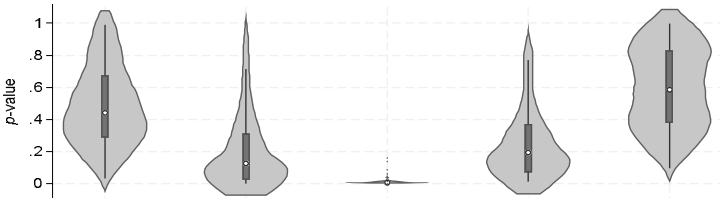
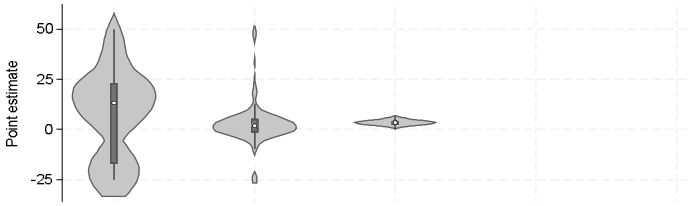
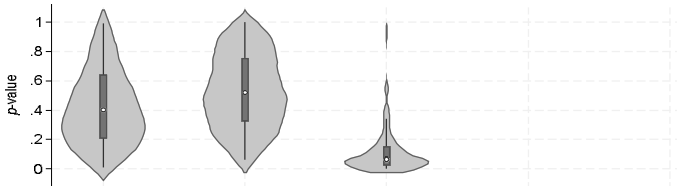
Estado do Rio Grande do Sul (2018). Genealogia dos municípios do Rio Grande do Sul. Secretaria de Planejamento, Governança e Gestão (SPGG). Departamento de Planejamento Governamental. Porto Alegre: SPGG.

Lipscomb, M., Mobarak, A. M., & Barham, T. (2013*b*). Replication data for: Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil. *American Economic Association* [publisher], Inter-university Consortium for Political and Social Research [distributor]. doi: 10.3886/E113850V1.

# Appendix B. Additional data information

**Table B1: Revised results on electrification effects with different replication procedures**

| | *Panel A. Results for median seed run* | | | | |
|---|---|---|---|---|---|
| | Step 3 with alternative Amazon | | Order change | No revised aggregation | Step 5 w. alternative outcome adjustment |
| | North | jungle | Step 5 excl. Step 4 | Step 5 | |
| | (1) | (2) | (3) | (4) | (5) |
| HDI (IV) | 0.16 | 0.32** | 0.12*** | 0.06 | 0.02 |
| | (0.10) | (0.16) | (0.04) | (0.05) | (0.04) |
| Housing values (IV) | 12.53** | 1.99 | 3.46** | – | – |
| | (6.39) | (3.32) | (1.37) | | |
| First-stage coefficient | 0.09** | 0.09** | 0.30*** | 0.30*** | 0.30*** |
| | (0.04) | (0.04) | (0.05) | (0.05) | (0.05) |
| First-stage *F*-Statistic | 4.31 | 4.86 | 31.96 | 31.96 | 31.96 |
| Number of observations | 8726 | 8726 | 8726 | 8726 | 8726 |
| IV simulation model | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced |
| Consistent Amazon definition | yes (**North**) | yes (**jungle**) | yes (biome) | yes (biome) | yes (biome) |
| Adjustment and aggregation corrections | no | no | no | **adjust-ment only** | **alternative adjustment** |
| 2SLS time-interaction term | Amazon | Amazon | **quartic of suitability** | quartic of suitability | quartic of suitability |

| | *Panel B. Results across 100 seed runs* | | | | |
|---|---|---|---|---|---|
| | Step 3 with alternative Amazon | | Order change | No revised aggregation | Step 5 w. alternative outcome adjustment |
| | North | jungle | Step 5 excl. Step 4 | Step 4 | |
| | (1) | (2) | (3) | (4) | (5) |
| *HDI* | | | | | |
| point estimate distribution |  | | | | |
| *p*-value distribution |  | | | | |
| share of *p*-values <0.1[a] | 4% | 48% | 98% | 33% | 1% |
| *Housing values* | | | | | |
| point estimate distribution |  | | | | |
| *p*-value distribution |  | | | | |
| share of *p*-values <0.1[a] | 6% | 3% | 67% | – | – |
| Share of runs with significant mechanisms[b] | | | | | |
| Education | n/c | n/c | ◕◔○○◕ | – | – |
| Employment | n/c | n/c | ◑◕○◕ | – | – |
| Health | n/c | n/c | ◑● | – | – |
| Income | n/c | n/c | ○○◔ | – | – |
| Population | n/c | n/c | ◕○○ | – | – |
| Share of first-stage *F*-Statistics <10 | 100% | 88% | 4% | 4% | 4% |
| IV simulation model | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced |
| Consistent Amazon definition | yes (**North**) | yes (**jungle**) | yes (biome) | yes (biome) | yes (biome) |
| Adjustment and aggregation corrections | no | no | no | **adjust-ment only** | **alternative adjustment** |
| 2SLS time-interaction term | Amazon | Amazon | **quartic of suitability** | quartic of suitability | quartic of suitability |

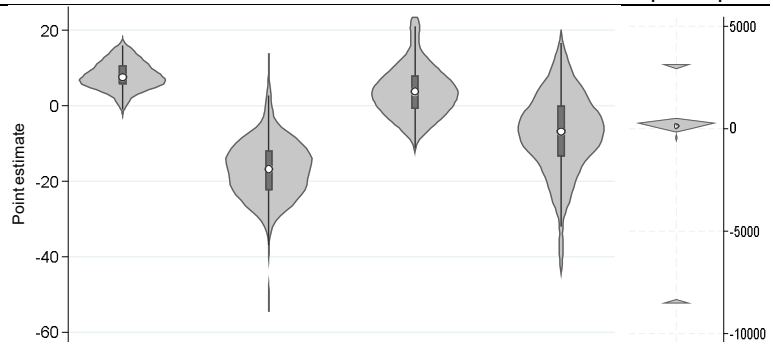
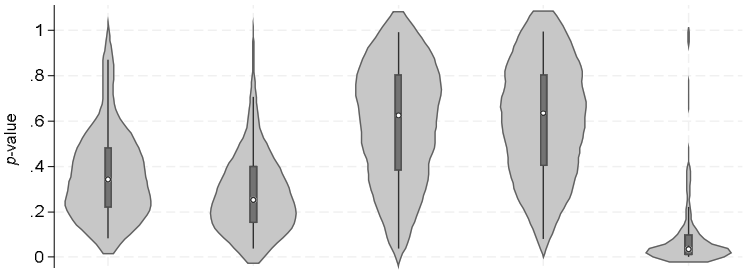
*Note: n/c = not calculated given the preponderance of weak first stages; – = not presented because no changes for Housing values and only partly changes for mechanisms. The change to the model in the respective step is highlighted*

*in bold. For expositional purposes, point estimates in the violin plots are top- and bottom-coded at 0.5 and -0.25 for HDI and 50 and -25 for housing values, respectively.* [a] *This share refers to p-values on estimates with a positive effect direction.* [b] *The full list of mechanisms is presented in Table B2. Significance refers to the 10% significance level of a two-sided t-test, again with a positive effect direction, for example a decrease for the poverty ratio. The differently filled circles reflect the shares of runs where the coefficients for the mechanisms were found significant at that level: ○: ≤10%; ◔: 10<x≤25%; ◑: 25<x≤50%; ◕: 50<x≤75%; ●: >75%. *** p<0.01, ** p<0.05, * p<0.1.*

**Table B2: List of mechanism variables studied in LMB**

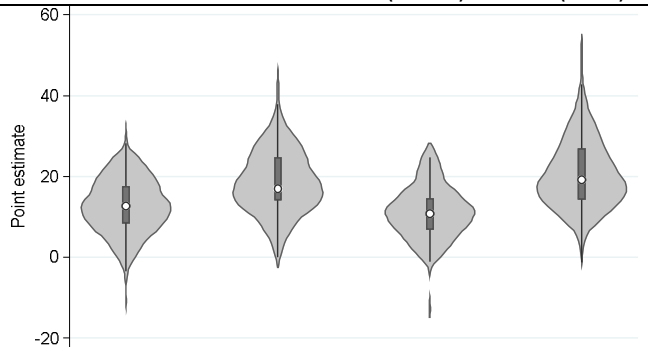
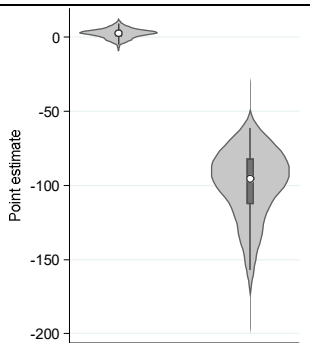
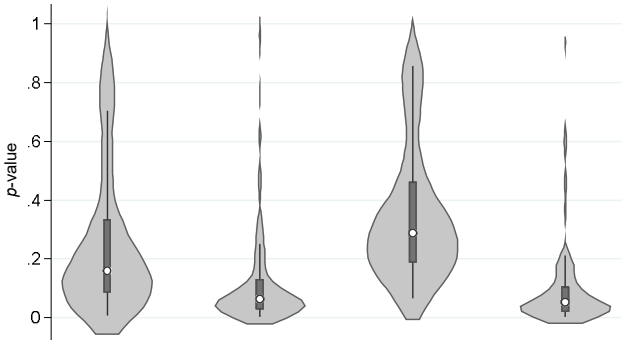
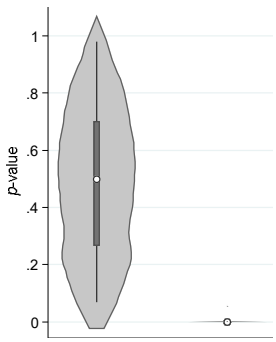
| |
|---|
| *Education* |
| HDI sub-component education |
| Illiteracy rate |
| Share of adults over age 15 with less than four years of formal education |
| Average years of Education |
| Human Capital per capita |
| *Employment* |
| Share of economically active population |
| Share of population formally employed |
| Urban share of population formally employed |
| Rural share of population formally employed |
| *Health* |
| HDI sub-component longevity |
| Infant mortality |
| *Income* |
| HDI sub-component income |
| Gross income per capita |
| Poverty ratio |
| *Population* |
| Share of in-migrants (past 5 years) |
| Population density |
| Share of urban population |

**Table B3: Revised results on mechanisms: education**

| | *Panel A. Results for median seed run on educational outcomes* | | | | |
|---|---|---|---|---|---|
| | HDI: Education | Illiteracy rate | Less than four years of education | Years of Education | Human Capital per capita |
| IV | 0.03<br>(0.03) | -5.78*<br>(3.44) | 0.75<br>(4.47) | -0.16<br>(0.37) | 20.53***<br>(7.70) |
| | *Panel B. Results across simulation model runs on educational outcomes* | | | | |
| | HDI: Education | Illiteracy rate | Less than four years of education | Years of Education | Human Capital per capita |
| point estimate distribution |  | | | | |
| *p*-value distribution |  | | | | |
| share of *p*-values <0.1[a] | 2% | 14% | 0% | 0% | 75% |
| IV simulation model | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced |
| Consistent Amazon def. | yes | yes | yes | yes | yes |
| Adjustment and aggregation corrections | yes | yes | yes | yes | yes |
| 2SLS time-interaction term | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability |

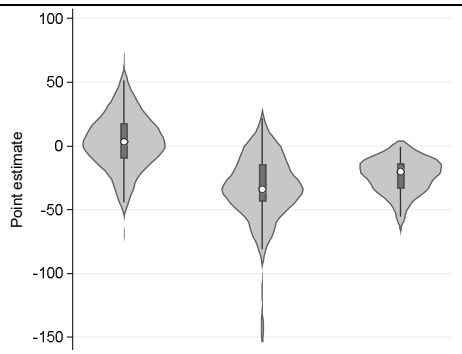
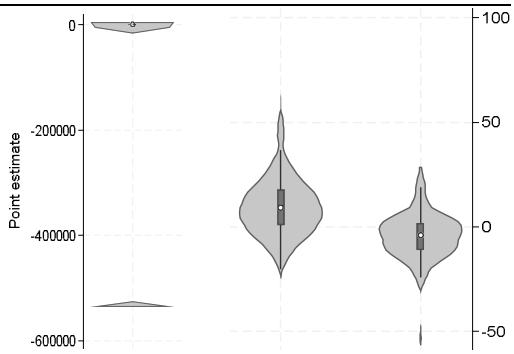
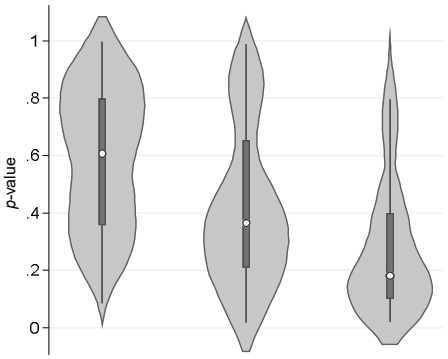
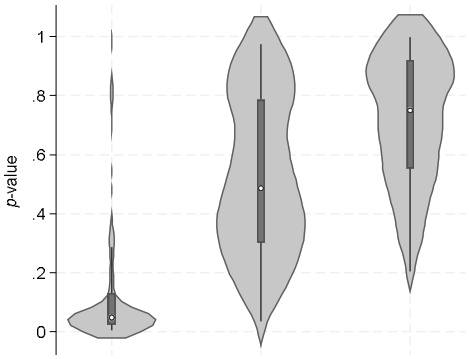
*Note: Point estimates expressed relative to the outcome mean.* [a] *This share refers to p-values on estimates with a positive effect direction. *** p<0.01, ** p<0.05, * p<0.1.*

**Table B4: Revised results on mechanisms: employment and health**

| | *Panel A. Results for median seed run* | | | | | |
|---|---|---|---|---|---|---|
| | *Employment outcomes* | | | | *Health outcomes* | |
| | Econo-mically active | Formal employment | Formal employment (urban) | Formal employment (rural) | HDI: Longevity | Infant mortality |
| IV | 0.05* (0.02) | 0.06** (0.02) | 0.03 (0.03) | 0.07*** (0.02) | 0.02 (0.02) | -63.04*** (13.13) |
| | *Panel B. Results across 100 seed runs* | | | | | |
| | *Employment outcomes* | | | | *Health outcomes* | |
| | Econo-mically active | Formal employment | Formal employment (urban) | Formal employment (rural) | HDI: Longevity | Infant mortality |
| point estimate distribution |  | | | |  | |
| *p*-value distribution |  | | | |  | |
| share of *p*-values <0.1[a] | 29% | 68% | 9% | 72% | 8% | 100% |
| IV simulation model | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced |
| Consistent Amazon def. | yes | yes | yes | yes | yes | yes |
| Adjustment and aggregation corrections | yes | yes | yes | yes | yes | yes |
| 2SLS time-interaction term | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability |

*Note: Point estimates expressed relative to the outcome mean.* [a] *This share refers to p-values on estimates with a positive effect direction. *** p<0.01, ** p<0.05, * p<0.1.*

**Table B5: Revised results on mechanisms: income and population**

| | *Panel A. Results for median seed run* | | | | | |
|---|---|---|---|---|---|---|
| | *Income outcomes* | | | *Population outcomes* | | |
| | HDI: Income | Gross income per capita | Poverty ratio | Share of in-migrants | Population density | Share of urban population |
| IV | 0.00<br>(0.09) | -0.03<br>(0.03) | -11.51<br>(7.92) | -10.29<br>(66.82) | 15.73<br>(10.24) | -0.01<br>(0.08) |
| | *Panel B. Results across 100 seed runs* | | | | | |
| | *Income outcomes* | | | *Population outcomes* | | |
| point estimate distribution |  | | |  | | |
| *p*-value distribution |  | | |  | | |
| share of *p*-values <0.1[a] | 0% | 0% | 25% | 73% | 1% | 0% |
| IV simulation model | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced | LMS model reproduced |
| Consistent Amazon def. | yes | yes | yes | yes | yes | yes |
| Adjustment and aggregation corrections | yes | yes | yes | yes | yes | yes |
| 2SLS time-interaction term | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability | quartic of suitability |

*Note: Point estimates expressed relative to the outcome mean.* [a] *This share refers to p-values on estimates with a positive effect direction. *** p<0.01, ** p<0.05, * p<0.1.*