Hidden in plain sight

Influential sets in linear regression

Nikolas Kuschnig[*], Gregor Zens, Jesús Crespo Cuaresma

IAAE 2023, Oslo, 28th of June, 2023

Vienna University of Economics and Business
[*]nikolas.kuschnig@wu.ac.at

The motivation

How sensitive are our inferences?

The motivation

How sensitive are our inferences?

What if our results depend on **one** observation?

The motivation

Ow sensitive are our inferencs?

What if our results depend on one observation?

- The issue has been studied in detail.

The motivation

*How sensitive are our **in**fe**re**nces?*

What if our results depend on **a few** observations?

- For single observations, the issue has been studied in detail.

The motivation

*How sensitive are our fen**c**es?*

What if our results depend on a few observations?

- For single observations, the issue has been studied in detail.
- The issue is not well understood, and quickly intractable.

The motivation

*How sensitive are our fe**n**ces?*

What if our results depend on a few observations?

- For single observations, the issue has been studied in detail.
- The issue is not well understood, and quickly intractable.

Consequences can be dire.

The setting

We investigate the sensitivity of inferences to **influential sets**.

> *A set of observations $\mathcal{S}$ is influential if **its omission** has a large impact on some measure of interest λ when compared to others.*

The setting

We investigate the sensitivity of inferences to **influential sets**.

> *A set of observations $\mathcal{S}$ is influential if **its omission** has a large impact on some measure of interest λ when compared to others.*

We want *sets with maximal influence* $\Delta(\mathcal{S})$ at a given size, to find the

- **minimal influential set** $\mathcal{S}^{**}$ — that is
- the *smallest set whose removal* ***overturns a result of interest***.

The setting

We investigate the sensitivity of inferences to **influential sets**.

> *A set of observations $\mathcal{S}$ is influential if **its omission** has a large impact on some measure of interest λ when compared to others.*

We want *sets with maximal influence* $\Delta(\mathcal{S})$ at a given size, to find the

- **minimal influential set** $\mathcal{S}^{**}$ — that is
- the *smallest set whose removal **overturns a result of interest***.

Example — 'The Blessing of Bad Geography in Africa'

'[...] the differential effect of ruggedness is statistically significant and economically meaningful, [...]' (Nunn and Puga, 2012)

Issue #1 — computation

Exactly determining the minimal influential set is usually **impossible**.

Issue #1 — computation

Exactly determining the minimal influential set is usually *impossible*.

1. There are $\binom{N}{N_\alpha}$ possible sets, where N_α is the set size, $|\mathcal{S}^{**}|$.
2. We need to compute λ, the quantity of interest, for each one.

Consider $N = 1,000$*, allowing for* $N_\alpha = 10$*, and assume that calculating* λ *takes one μs. Your sensitivity check will take about 8.35 billion years.*

Issue #1 — computation

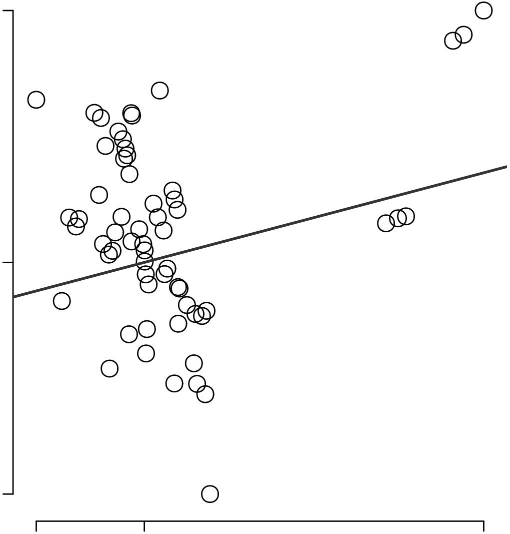
Exactly determining the minimal influential set is usually *impossible*.

1. There are $\binom{N}{N_\alpha}$ possible sets, where N_α is the set size, $|\mathcal{S}^{**}|$.
2. We need to compute λ, the quantity of interest, for each one.

Consider $N = 1,000$, allowing for $N_\alpha = 10$, and assume that calculating λ takes one µs. Your sensitivity check will take about 8.35 billion years.

There is a number of useful results to quickly evaluate λ and $\Delta(\mathcal{S})$, but we need to **approximate the set** in all but the simplest cases.

Issue #2 — joint influence

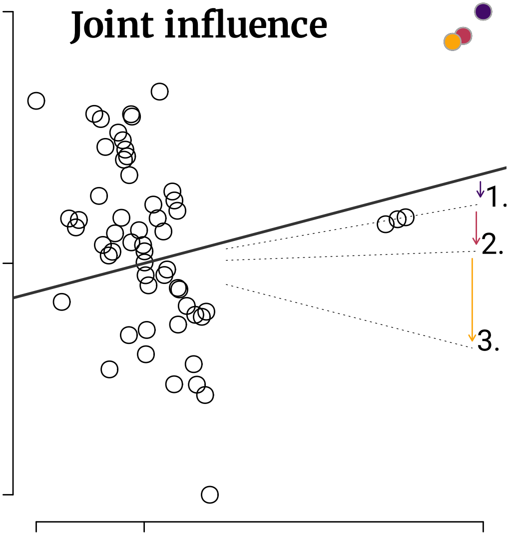


Consider the model $\mathbf{y} = \mathbf{x}\beta + \varepsilon$, with

$$\lambda(\mathcal{S}) = \left(\mathbf{x}'_{(\mathcal{S})}\mathbf{x}_{(\mathcal{S})}\right)^{-1} \mathbf{x}'_{(\mathcal{S})}\mathbf{y}_{(\mathcal{S})},$$

where $\mathcal{S}$ is a set of observations, and subscripts indicate removal.

Issue #2 — joint influence



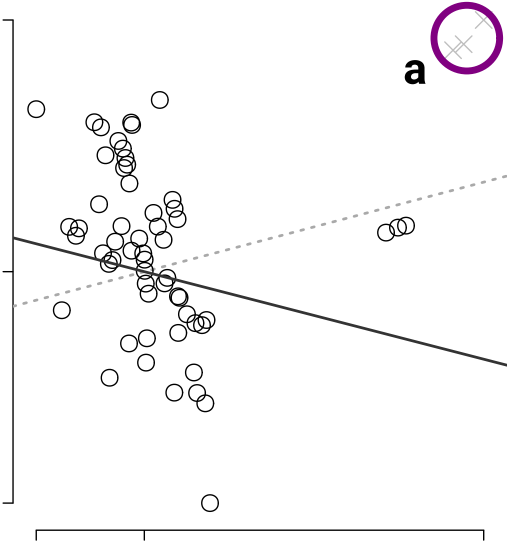
Consider the model $\mathbf{y} = \mathbf{x}\beta + \varepsilon$, with

$$\lambda(\mathcal{S}) = \left(\mathbf{x}'_{(\mathcal{S})}\mathbf{x}_{(\mathcal{S})}\right)^{-1} \mathbf{x}'_{(\mathcal{S})}\mathbf{y}_{(\mathcal{S})},$$

where $\mathcal{S}$ is a set of observations, and subscripts indicate removal.

- The **influence of a set** may
- exceed the individual (full-sample) influences of its members —
- sets may be **jointly influential**.

Issue #3 — masking

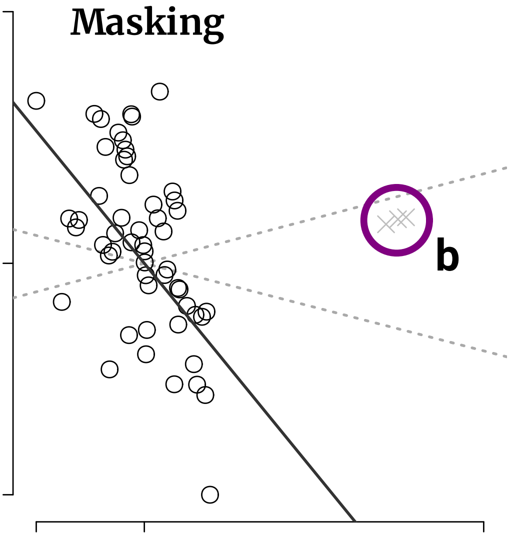


Consider the model $\mathbf{y} = \mathbf{x}\beta + \varepsilon$, with

$$\lambda(\mathcal{S}) = \left(\mathbf{x}'_{(\mathcal{S})}\mathbf{x}_{(\mathcal{S})}\right)^{-1}\mathbf{x}'_{(\mathcal{S})}\mathbf{y}_{(\mathcal{S})},$$

where $\mathcal{S}$ is a set of observations, and subscripts indicate removal.

- The set marked **'a'** is **highly influential** on the slope.



Consider the model $\mathbf{y} = x\beta + \varepsilon$, with

$$\lambda(\mathcal{S}) = \left(\mathbf{x}'_{(\mathcal{S})}\mathbf{x}_{(\mathcal{S})}\right)^{-1} \mathbf{x}'_{(\mathcal{S})}\mathbf{y}_{(\mathcal{S})},$$

where $\mathcal{S}$ is a set of observations, and subscripts indicate removal.

- The set marked **‘a’** is **highly influential** on the slope.
- However, it initially **masks** the influential set marked **‘b’**.

Identifying influential sets

How to identify a minimal influential set?

How to identify a minimal influential set?

We consider *three algorithms* to approximate $\mathcal{S}$ and $\Delta(\hat{\mathcal{S}})$, that are

- easy to implement,
- computationally tractable,
- differently trade speed for accuracy.

How to identify a minimal influential set?

We consider *three algorithms* to approximate $\mathcal{S}$ and $\Delta(\hat{\mathcal{S}})$, that are

- easy to implement,
- computationally tractable,
- differently trade speed for accuracy.

We focus on the most accurate and precise one — an **adaptive search**.[1]

[1] The others use ▸ A0 the *full-sample influence* (akin to the approach by Broderick, Giordano, and Meager, 2023), and ▸ A1 a *binary search* for improved speed.

The algorithms — an adaptive approximation

Algorithm 2

Idea: Greedily build approximations to $\mathcal{S}$.

The algorithms — an adaptive approximation

Algorithm 2

Idea: Greedily build approximations to $\mathcal{S}$.

0. Let $\hat{\mathcal{S}} \leftarrow \emptyset$.

The algorithms — an adaptive approximation

Algorithm 2

Idea: Greedily build approximations to $\mathcal{S}$.

0. Let $\hat{\mathcal{S}} \leftarrow \varnothing$.
1. Compute $\Delta(\hat{\mathcal{S}} \cup \{j\})$ for each $j \notin \hat{\mathcal{S}}$.

The algorithms — an adaptive approximation

Algorithm 2

Idea: Greedily build approximations to $\mathcal{S}$.

0. Let $\hat{\mathcal{S}} \leftarrow \varnothing$.
1. Compute $\Delta(\hat{\mathcal{S}} \cup \{j\})$ for each $j \notin \hat{\mathcal{S}}$.
2. Let $\hat{\mathcal{S}} \leftarrow \hat{\mathcal{S}} \cup \arg\max \Delta(\hat{\mathcal{S}} \cup \{j\})$.

The algorithms — an adaptive approximation

Algorithm 2

Idea: Greedily build approximations to $\mathcal{S}$.

0. Let $\hat{\mathcal{S}} \leftarrow \varnothing$.
1. Compute $\Delta(\hat{\mathcal{S}} \cup \{j\})$ for each $j \notin \hat{\mathcal{S}}$.
2. Let $\hat{\mathcal{S}} \leftarrow \hat{\mathcal{S}} \cup \arg\max \Delta(\hat{\mathcal{S}} \cup \{j\})$.
3. Go to step 1, unless $\Delta(\hat{\mathcal{S}}) > \Delta^*$ or $|\hat{\mathcal{S}}| > U$.

The algorithms — an adaptive approximation

Algorithm 2

Idea: Greedily build approximations to $\mathcal{S}$.

0. Let $\hat{\mathcal{S}} \leftarrow \varnothing$.
1. Compute $\Delta(\hat{\mathcal{S}} \cup \{j\})$ for each $j \notin \hat{\mathcal{S}}$.
2. Let $\hat{\mathcal{S}} \leftarrow \hat{\mathcal{S}} \cup \arg\max \Delta(\hat{\mathcal{S}} \cup \{j\})$.
3. Go to step 1, unless $\Delta(\hat{\mathcal{S}}) > \Delta^*$ or $|\hat{\mathcal{S}}| > U$.

This way, we can *adapt for masking* at $\mathcal{O}(N_\alpha)$ complexity. Computing Δ dominates, but *updating formulae* and *approximations* allow for computationally efficient implementation.

The influence and computing Δ

Example — 'The Blessing of Bad Geography in Africa'

Rugged terrain hinders development globally. Nunn and Puga find a *different* **statistically and economically significant** *effect* in Africa.

The influence and computing Δ

Example — 'The Blessing of Bad Geography in Africa'

Rugged terrain hinders development globally. Nunn and Puga find a *different* **statistically and economically significant** *effect* in Africa.

In most regression analyses, we tend to care about the

- estimated **coefficient** ($\hat{\beta}$), and
- **uncertainty** around it (standard errors or t values).

The influence and computing Δ

Example — 'The Blessing of Bad Geography in Africa'

Rugged terrain hinders development globally. Nunn and Puga find a *different* **statistically and economically significant** *effect* in Africa.

In most regression analyses, we tend to care about the

- estimated **coefficient** ($\hat{\beta}$), and
- **uncertainty** around it (standard errors or t values).

For these, we have closed form results and efficient updating formulae, e.g.

$$\Delta(\{i\}) = \beta_{(\emptyset)} - \beta_{(\{i\})} = \frac{(\mathbf{X}'\mathbf{X})^{-1} x_i' e_i}{1 - h_i}.$$

A demonstration

*What does a **minimal influential set** look like in practice?*

A demonstration

*What does a **minimal influential set** look like in practice?*

1. First, we'll have a look at the *univariate regression* from earlier.

A demonstration

What does a ***minimal influential set*** *look like in practice?*

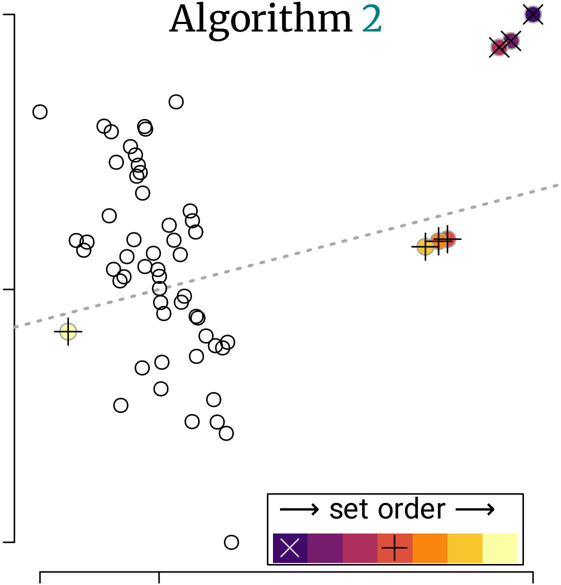
1. First, we'll have a look at the *univariate regression* from earlier.
2. Then, we'll investigate three papers on **long-term development**, on

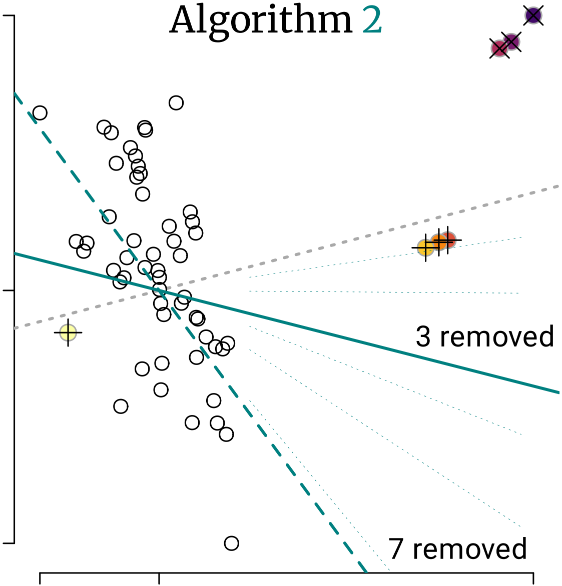
A demonstration

*What does a **minimal influential set** look like in practice?*

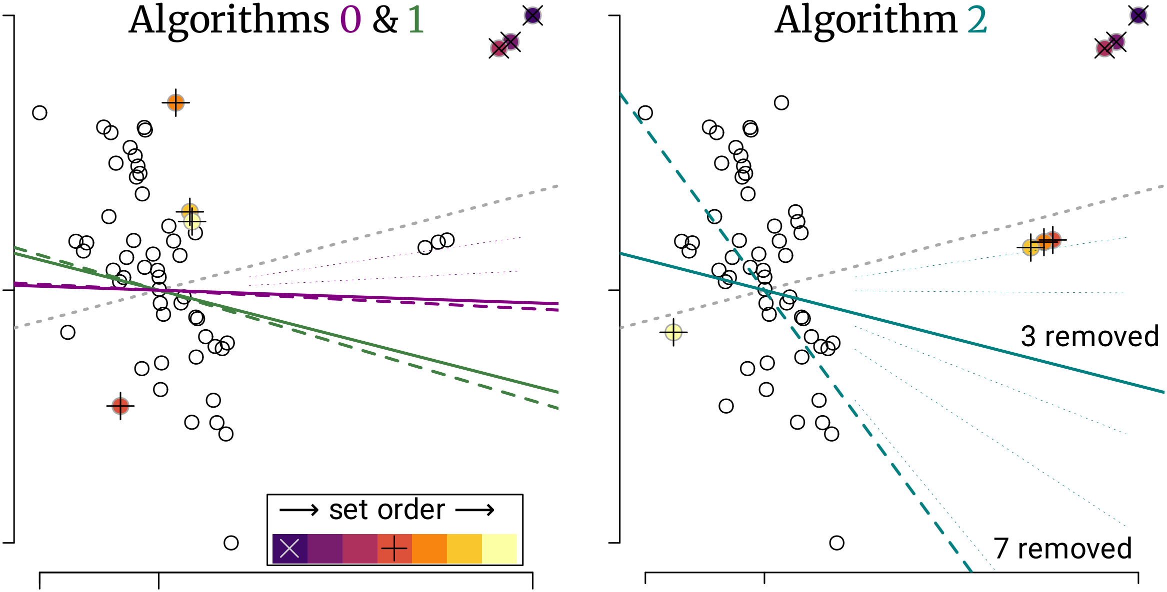
1. First, we'll have a look at the *univariate regression* from earlier.
2. Then, we'll investigate three papers on **long-term development**, on
   - the blessing of **bad geography** in Africa (Nunn and Puga, 2012),
   - the slave trades and **mistrust** (Nunn and Wantchekon, 2011), and
   - the effect of the **Tsetse fly** (Alsan, 2015).

Algorithm 2





Algorithm 2
3 removed
7 removed


Algorithms 0 & 1
→ set order →
Algorithm 2
3 removed
7 removed

Applications — influential sets and ruggedness

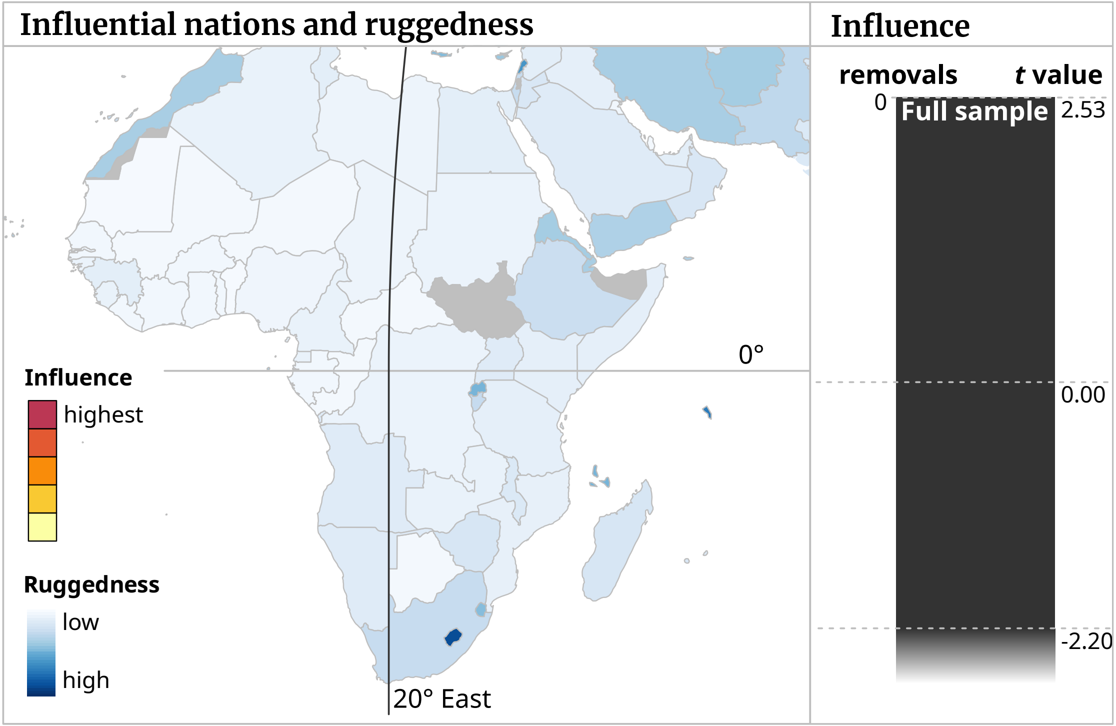
log GDP/capita ~	Baseline	Plain
ruggedness, Africa†	0.321	0.302
	(2.53)	(2.32)
ruggedness	-0.231	-0.193
	(-2.99)	(-2.38)
coast distance	Yes	Yes
other controls	Yes	–
observations	170	170

The (t values) are based on HC1 standard errors. The 'thresholds' indicate the number of removed observation that nullify significance (at the 5% level), [flip the sign], and {significantly flip the sign}.

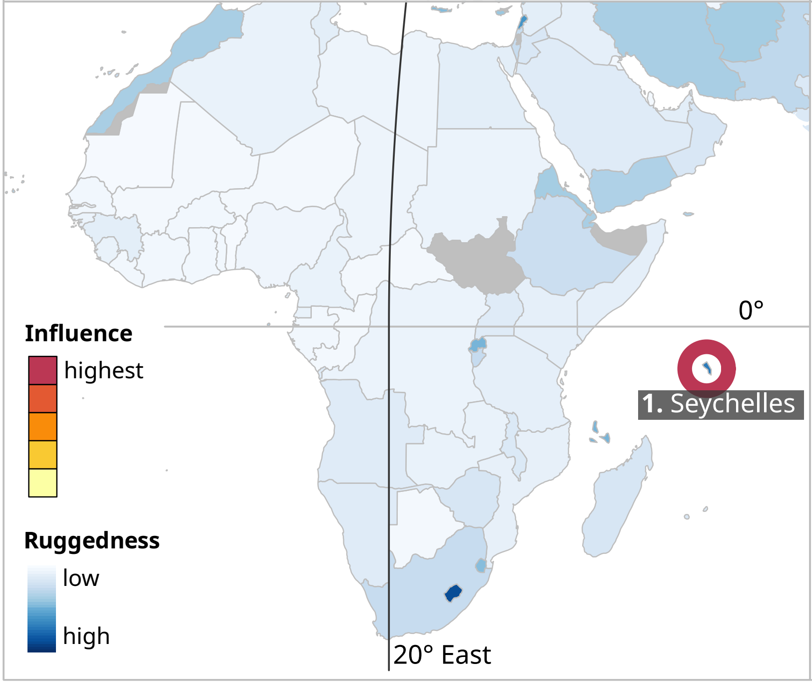
Applications — influential sets and ruggedness

log GDP/capita ~	Baseline	Plain
ruggedness, Africa†	0.321	0.302
	(2.53)	(2.32)
ruggedness	-0.231	-0.193
	(-2.99)	(-2.38)
coast distance	Yes	Yes
other controls	Yes	–
observations	170	170
thresholds†	2 [5]{11}	2[7]{16}

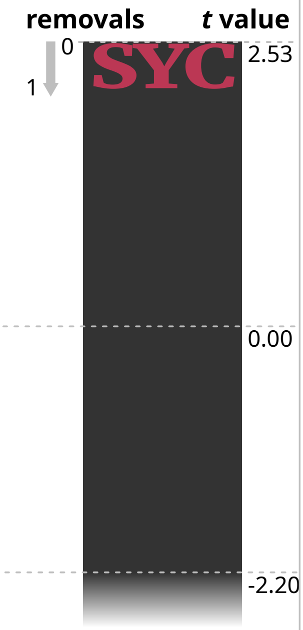
The (t values) are based on HC1 standard errors. The 'thresholds' indicate the number of removed observation that nullify significance (at the 5% level), [flip the sign], and {significantly flip the sign}.


Influential nations and ruggedness
Influence
0°
20° East
Influence
highest
Ruggedness
low
high
removals
t value
0
Full sample
2.53
0.00
-2.20

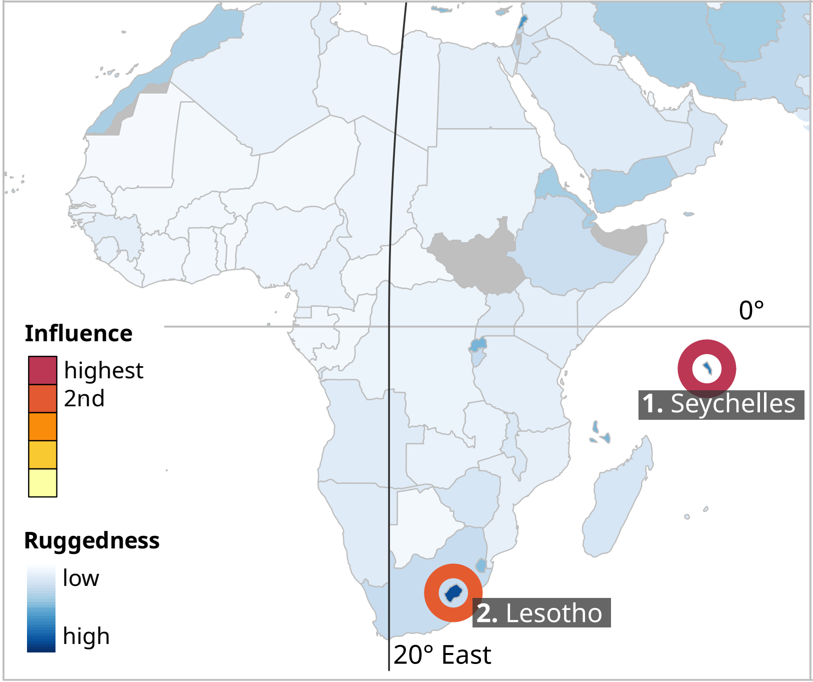
Influential nations and ruggedness



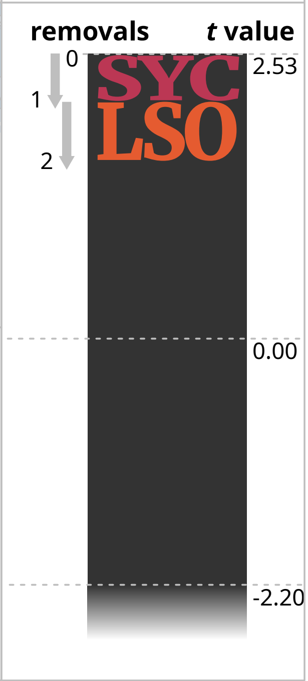
Influence



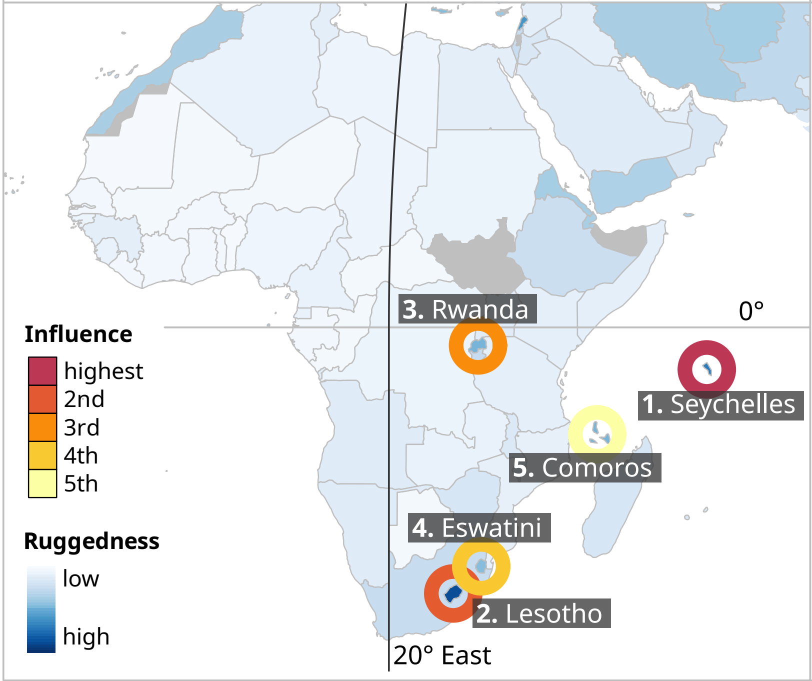
Influential nations and ruggedness



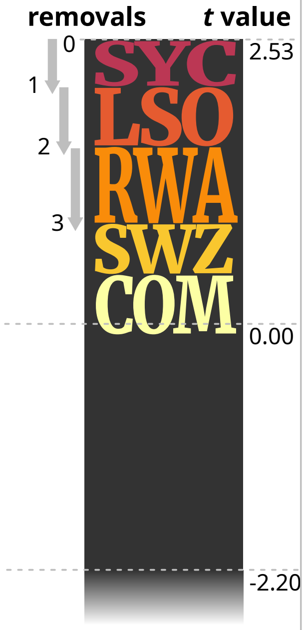
Influence



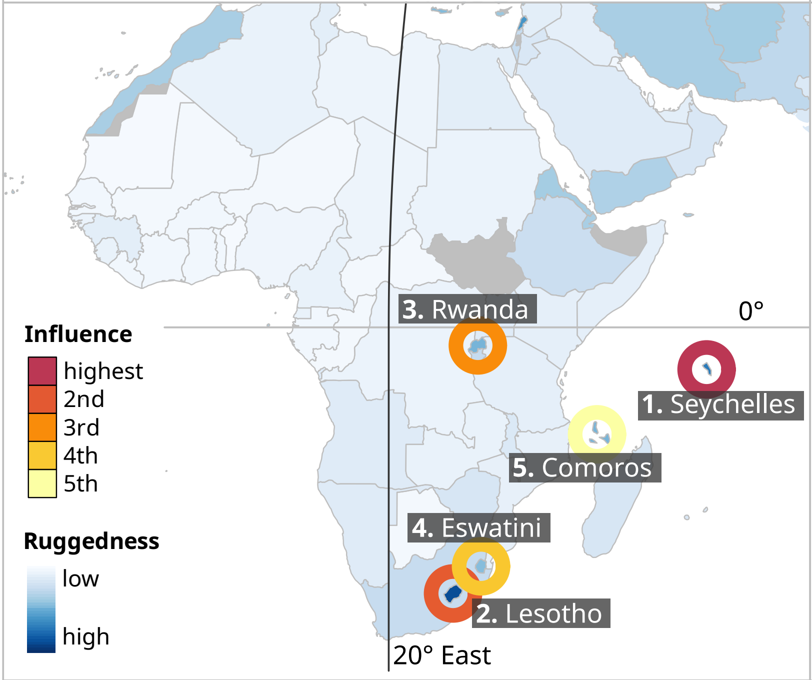
Influential nations and ruggedness



Influence



Influential nations and ruggedness



Influence

Applications — effects of the Tsetse fly

	Animals	Intensive	Plow	Female	Density	Slavery	Centralized
TSI†	-0.231	-0.09	-0.057	0.206	-0.745	0.101	-0.075
	(-5.47)	(-3.29)	(-2.54)	(3.41)	(-3.25)	(2.51)	(-2.12)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
M-robust	Yes	Yes	No	Yes	Yes	No	Yes
S-robust	No	No	No	No	Yes	No	No
observations	484	485	484	315	398	446	467

The (t values) are based on clustered standard errors. Reported are the effects of the Tsetse suitability index (TSI) on — whether a precolonial ethnic group (1) possessed large domesticated 'Animals', (2) adopted 'Intensive' agriculture, (3) adopted the 'Plow', (4) had 'Female' participation in agriculture, (5) log population 'Density', (6) practiced indigenous 'Slavery', and (7) had a 'Centralized' state.

Applications — effects of the Tsetse fly

	Animals	Intensive	Plow	Female	Density	Slavery	Centralized
TSI†	-0.231	-0.09	-0.057	0.206	-0.745	0.101	-0.075
	(-5.47)	(-3.29)	(-2.54)	(3.41)	(-3.25)	(2.51)	(-2.12)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
M-robust	Yes	Yes	No	Yes	Yes	No	Yes
S-robust	No	No	No	No	Yes	No	No
observations	484	485	484 (37)	315	398	446	467
thresholds†	33[58]{79}	7[25]{41}	3 [12]{17}	12[30]{48}	9 [27]{42}	4[22]{35}	1[16]{30}

The (t values) are based on clustered standard errors. Reported are the effects of the Tsetse suitability index (TSI) on — whether a precolonial ethnic group (1) possessed large domesticated 'Animals', (2) adopted 'Intensive' agriculture, (3) adopted the 'Plow', (4) had 'Female' participation in agriculture, (5) log population 'Density', (6) practiced indigenous 'Slavery', and (7) had a 'Centralized' state.

Interpretation

Influential sets **can provide contextual insights**, but they **cannot** serve as a *conclusive robustness check* on their own.

Interpretation

Influential sets **can provide contextual insights**, but they **cannot** serve as a *conclusive robustness check* on their own.

If results *seem sensitive*, …

- we may be searching for the **needle in the haystack**,
  - + We should expect a small set *in relative terms*,
  - – but one with low cardinality indicates low power.

Interpretation

Influential sets **can provide contextual insights**, but they **cannot** serve as a *conclusive robustness check* on their own.

If results *seem sensitive*, …

- we may be searching for the **needle in the haystack**,
  - \+ We should expect a small set *in relative terms*,
  - – but one with low cardinality indicates low power.
- or there should be plenty of needles.
  - ! We have an *outlier problem*, and some **data to investigate** —
  - ? there may be confounders, heterogeneous effects, etc.

Summary

To wrap up — we were looking for **minimal influential sets**, an

Summary

To wrap up — we were looking for **minimal influential sets**, an

- **intuitive** (*two nations remove significance*),
- **insightful** (*confounders, heterogeneity, validity*), and
- **widely applicable** (*size, clustered errors, 2SLS*) sensitivity check.

Summary

To wrap up — we were looking for **minimal influential sets**, an

- intuitive (*two nations remove significance*),
- insightful (*confounders, heterogeneity, validity*), and
- widely applicable (*size, clustered errors, 2SLS*) sensitivity check.

We've also caused some issues, e.g.

- How to find **better sets faster**?
- How **problematic is influence**? ▸ See more

Summary

To wrap up — we were looking for **minimal influential sets**, an

- intuitive (*two nations remove significance*),
- insightful (*confounders, heterogeneity, validity*), and
- widely applicable (*size, clustered errors, 2SLS*) sensitivity check.

We've also caused some issues, e.g.

- How to find better sets faster?
- How problematic is influence? ‣ See more

Find the paper, an **R package**, and an **interactive illustration** online.

References i

Marcella Alsan.
The effect of the TseTse fly on African development.
American Economic Review, 105(1):382–410, 2015.

Tamara Broderick, Ryan Giordano, and Rachael Meager.
An automatic finite-sample robustness metric: can dropping a little data change conclusions?, 2020.

Felipe Valencia Caicedo.
The Mission: human capital transmission, economic persistence, and culture in South America.
Quarterly Journal of Economics, 134(1):507–556, 2019.

Jesús Crespo Cuaresma, Stephan Klasen, and Konstantin M. Wacker.
When do we see poverty convergence?
Oxford Bulletin of Economics and Statistics, 2022.

Mauricio Drelichman, Jordi Vidal-Robert, and Hans-Joachim Voth.
The long-run effects of religious persecution: Evidence from the spanish inquisition.
Proceedings of the National Academy of Sciences, 118(33):e2022881118, 2021.

Nathan Nunn.
The historical roots of economic development.
Science, 367(6485):eaaz9986, 2020.

Nathan Nunn and Diego Puga.
Ruggedness: the blessing of bad geography in Africa.
Review of Economics and Statistics, 94(1):20–36, 2012.

Nathan Nunn and Leonard Wantchekon.
The slave trade and the origins of mistrust in Africa.
American Economic Review, 101(7):3221–52, 2011.

The algorithms — an initial approximation

Algorithm 0

Idea: Approximate $\mathcal{S}$ based on initial influence and Δ via summation.

0. Compute $\Delta(\{i\})$ for each observation i, let $\hat{\mathcal{S}} \leftarrow \emptyset$.
1. Let $\hat{\mathcal{S}} \leftarrow \hat{\mathcal{S}} \cup \arg\max \Delta(\{j\})$, for $j \notin \hat{\mathcal{S}}$.
2. Let $\hat{\Delta}(\hat{\mathcal{S}}) \leftarrow \sum \Delta(\{k\})$ for all $k \in \hat{\mathcal{S}}$.
3. Go to step 1, unless $\hat{\Delta} > \Delta^*$ or $|\hat{\mathcal{S}}| > U$.

At $\mathcal{O}(1)$ complexity, **computing Δ dominates**. Broderick, Giordano, and Meager (2020) use a similar approach, approximating Δ Details. Back

The algorithms — divide and conquer

Algorithm 1

Idea: Approximate $\mathcal{S}$ based on initial influence; binary-search for Δ^*.

1. Compute $\Delta(\{i\})$ for each observation i.
2. Create the ordered set $\mathcal{T}$ by ranking $\Delta(\{i\})$.
3. Binary-search for the smallest Δ^* in the interval (L, U).
   - Let $\hat{\mathcal{S}}$ be the first $(L + U)/2$ elements of $\mathcal{T}$.
   - Compute $\Delta(\hat{\mathcal{S}})$.
   - Adapt the lower or upper bound until done.

This adaptation yields improved precision at $\mathcal{O}(\log U)$ complexity. Back

Broderick, Giordano and Meager (2020)

'Can Dropping a Little Data Change Conclusions?' — the authors check using the 'Approximate Maximum Influence Perturbation' (AMIP).

- Computation of AMIP is effectively instant.
  - In our setting, their algorithm is a special case of Algorithm 0.
  - They use a linear approximation to compute Δ.
- Accuracy suffers, **especially when influential sets are present**.
  - There are *masking* issues and *downward bias*, akin to Algorithm 0.
  - The AMIP approximation of $\beta_{(\emptyset)} - \beta_{(\{i\})}$ *discards the leverage*, whereas

$$\text{influence} = f(\text{errors}, \text{leverage}).$$

- As a result, there is a high risk of *false negatives*.

Microcredit — seven randomised control trials

Sensitivity of the average treatment effect of microcredits

study region	BIH		MON		ETH		MEX		MOR		PHI		IND	
algorithm	(0)	(2)	(0)	(2)	(0)	(2)	(0)	(2)	(0)	(2)	(0)	(2)	(0)	(2)
sign-switch	14	13	16	15	1	1	1	1	11	11	9	9	6	6
significance	49	39	43	37	117	13	20	12	35	33	74	54	41	35
observations	1,195		961		3,113		16,560		5,498		1,113		6,863	

The reported values are the number of removals needed to induce a sign-switch of the average treatment effect, and have this sign-flipped coefficient become significant (at the 1% level) using Algorithm 0 and 2. Algorithm 2 outperforms consistently, but few observations are needed to overturn results in all cases.

Go back

Learning from influential sets — ruggedness

$\log$ GDP/capita $\sim$	Baseline	Plain	Robust-M	Population	Area
ruggedness, Africa†	0.321	0.302	0.325	0.190	0.215
	(2.53)	(2.32)	(2.46)	(1.66)	(1.63)
ruggedness	-0.231	-0.193	-0.251	-0.231	-0.238
	(-2.99)	(-2.38)	(-3.23)	(-2.94)	(-3.08)
coast distance	Yes	Yes	Yes	Yes	Yes
population in 1400	–	–	–	Yes	–
land area	–	–	–	–	Yes
other controls	Yes	–	Yes	Yes	Yes
observations	170	170	170	168	170
thresholds†	2[5]{11}	2[7]{16}	–	–[3]{6}	–[4]{8}

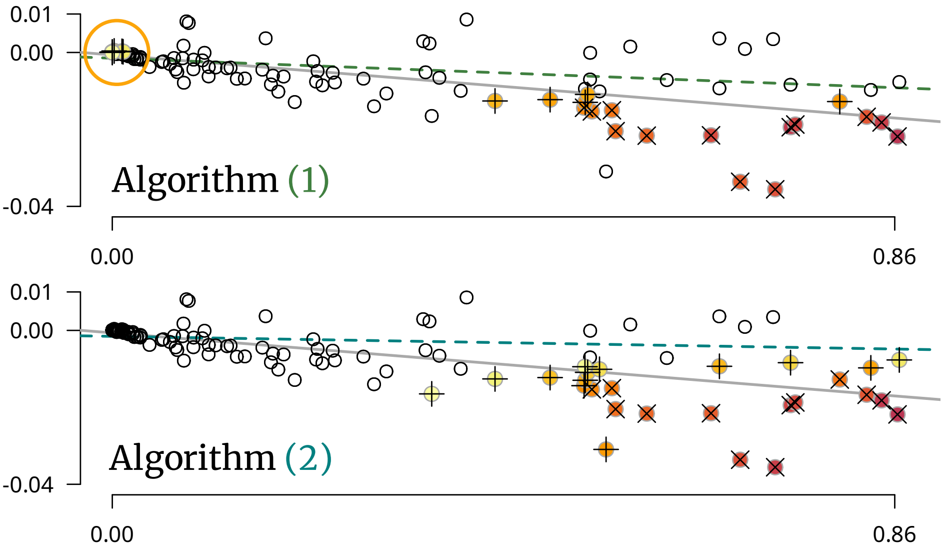
The 'thresholds' indicate the number of removed observation that nullify significance (at the 5% level), [flip the sign], and {significantly flip the sign}. The t values in (brackets) are based on HC1 errors. ▸ Go back

The origins of mistrust

	Trust of relatives ~		Trust of neighbours ~	
	Pooled	West|East	Pooled	West|East
exports/area†	-0.133	-0.145	-0.159	-0.168
	(-3.68)	(-3.84)	(-4.67)	(-4.48)
exports/area, East		0.053		0.023
		(0.96)		(0.32)
individual controls	Yes	Yes	Yes	Yes
district controls	Yes	Yes	Yes	Yes
country fixed effects	Yes	Yes	Yes	Yes
observations	20,062	7,549 | 12,513	20,027	7,523 | 12,504
thresholds†	105[380]{656}	78[301]{532}	161[425]{768}	133[323]{527}
ethnicity clusters	185	62 | 123	185	62 | 123
district clusters	1,257	628 | 651	1,257	628 | 651

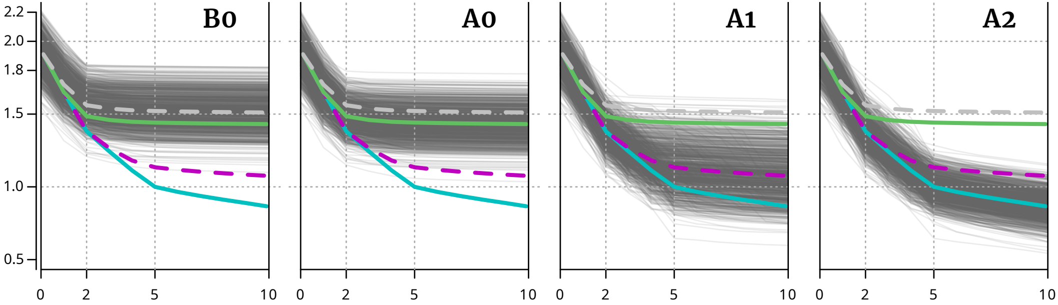
The (t values) are based on 2-way clustered standard errors. The 'thresholds' indicate the number of removed observation that nullify significance (1% level), [flip the sign], and {significantly do so}. Go back

Poverty convergence



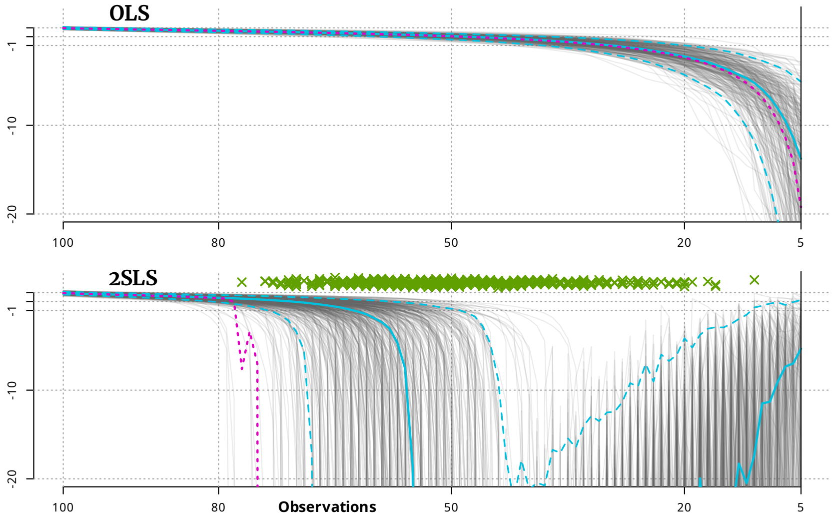
Data and regression line for the poverty convergence regression of Crespo Cuaresma et al. (2022), before (solid line) and after (dashed line) removing the influential set $\hat{\mathcal{S}}^*_{26}$. There are 126 observations in total.

Simulation results — algorithms



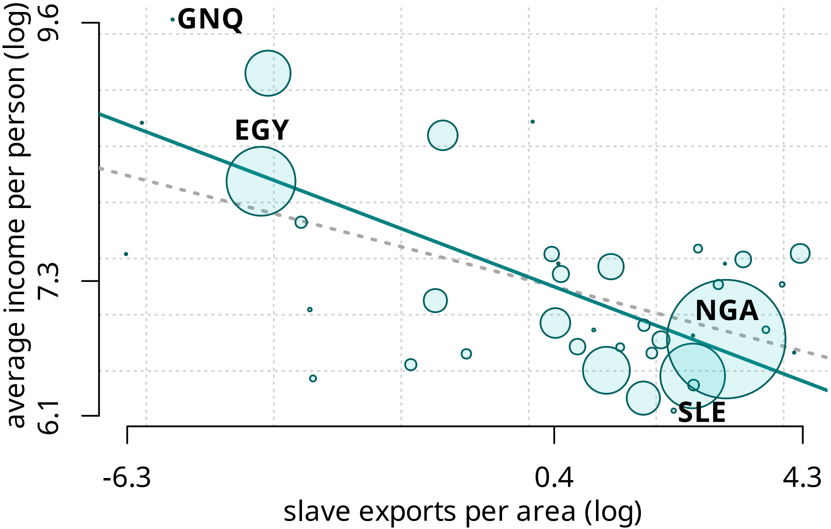
Transparent lines indicate individual runs, thick lines the average results of (from top to bottom) approach ‘B0’ (gray, dashed), ‘A0’ (green, solid), ‘A1’ (purple, dashed), and ‘A2’ (teal, solid). The vertical axis indicates estimates, the horizontal one the number of removals.

Simulation results — OLS and 2SLS



Transparent lines indicate individual simulations, thick ones the median (solid, blue), the 95% and 5% quantile (dashed), and the average (dotted, pink) of the estimate. Crosses at the top of the 2SLS panel indicate drop-outs due to pathological numerical stability (within machine precision).

Influence in $\frac{\text{outcome}}{\text{capita}}$ regressions



Average income in 2000 versus the past slave export density (following Nunn, 2020). Observations are weighted with their populations in 2000; lines indicate the weighted and unweighted (dashed) LS fit.

Do we expect large impacts of the Spanish inquisition?

Table 1: Inquisitorial Intensity on Modern Outcomes

	GDP/capita		Religiosity		Education		Trust	
	(LS)	(WLS)	(LS)	(WLS)	(LS)	(WLS)	(LS)	(WLS)
β	**-0.3962**	**-0.1870**	**0.4451**	0.1013	**-0.0535**	-0.0142	**-0.4003**	**-0.2180**
	(-9.582)	(-3.992)	(4.829)	(1.415)	(-2.333)	(-0.663)	(-2.803)	(-2.875)
$\boldsymbol{\theta}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{\mu}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2214	2214	2191	2191	2215	2215	976	976
R^2	0.491	0.569	0.429	0.548	0.572	0.635	0.05	0.074

Drelichman et al. (2021) investigate the long-run effects of religious persecution by the Spanish inquisition.

Confounded by influence?

Table 2: Missionaries on Modern Literacy

	Literacy			
	(LS)	(WLS)	(LS)	(WLS)
β	**0.0105**	0.0012	**0.0112**	-0.0010
	(2.860)	(0.208)	(2.261)	(-0.163)
$\boldsymbol{\theta}$	No	No	Yes	Yes
$\boldsymbol{\mu}$	Yes	Yes	Yes	Yes
N	549	549	548	548
R^2	0.042	0.082	0.073	0.172

Caicedo (2019) investigates the literacy impacts of Jesuit missions in South America.

Table 3: Cultural Punishment

	Income per person			
	(LS)	(WLS)	(LS)	(WLS)
β	**20.8954**	10.7928	**11.0059**	5.4548
	(6.087)	(1.928)	(2.999)	(1.001)
$\boldsymbol{\theta}$	No	No	Yes	Yes
$\boldsymbol{\mu}$	Yes	Yes	Yes	Yes
N	160	159	160	159
R^2	0.394	0.364	0.570	0.625

Michalopoulos and Xue (2021) investigate the economic impacts of punishment in oral traditions.