Yale

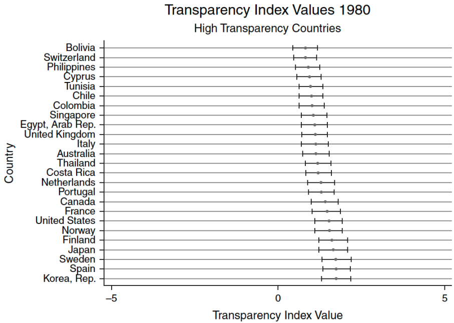
Causal Inference for IR and IPE with Substantive Applications

Carlos Felipe Balcazar

MacMillan Center

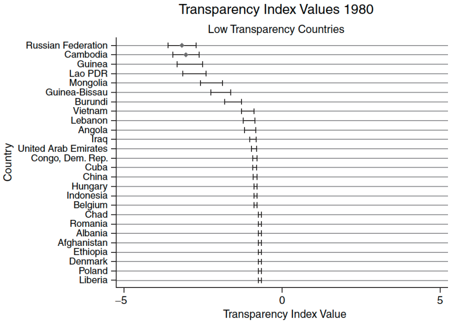
April, 2024

Transparency and cooperation



- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.

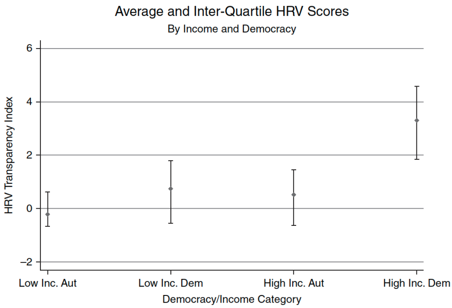
Transparency and cooperation



- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.

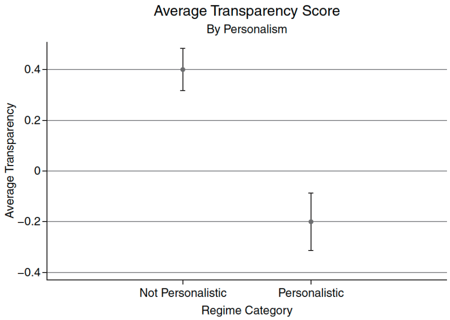
Yale

Transparency and cooperation



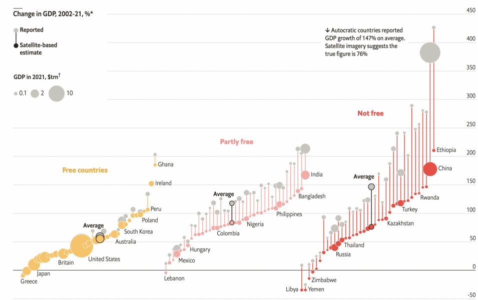
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.

Transparency and cooperation



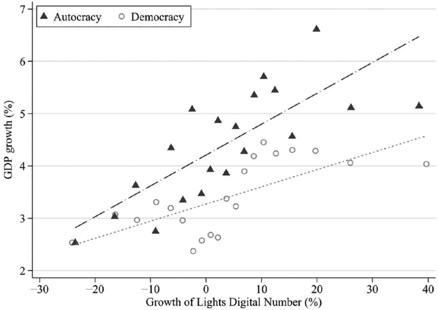
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.

Transparency and cooperation



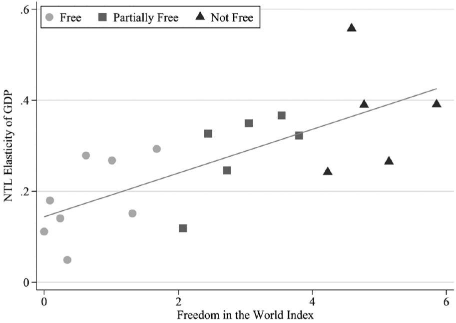
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.

Transparency and cooperation



- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.

Transparency and cooperation



- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.

Yale

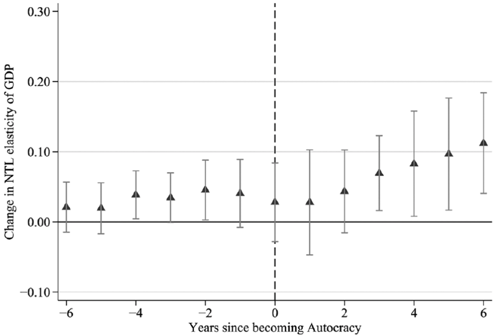
Transparency and cooperation

THE AUTOCRACY GRADIENT IN THE NTL ELASTICITY OF GDP SUBCOMPONENTS

	Consumption (1)	Investment (2)	Government (3)	Exports (4)	Imports (5)
$\ln(\mathrm{NTL})_{i,t}$	.184***	.353***	.210***	.354***	.253***
	[.041]	[.083]	[.060]	[.077]	[.054]
$\mathrm{FiW}_{i,t}$	−.003	.023	−.002	−.007	−.006
	[.035]	[.062]	[.041]	[.058]	[.042]
$\mathrm{FiW}_{i,t}^2$	−.002	−.010	−.001	−.004	−.005
	[.006]	[.012]	[.007]	[.011]	[.008]
$\ln(\mathrm{NTL})_{i,t} \times \mathrm{FiW}_{i,t}$	.004	.040***	.030***	.011	.013*
	[.006]	[.010]	[.007]	[.012]	[.008]

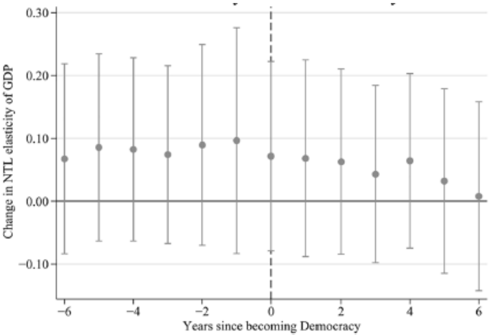
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.

Transparency and cooperation



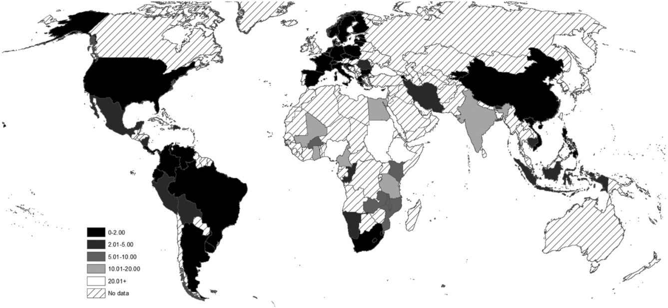
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.

Transparency and cooperation



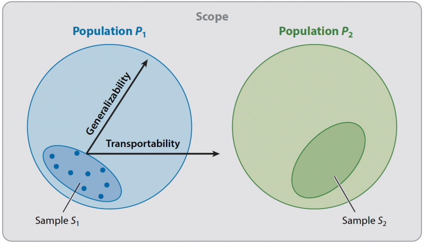
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.

Transparency and cooperation



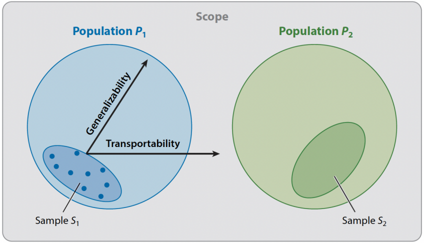
- Autocratic countries don't always report (accurate) information.
  - GDP, human rights abuses, corruption, health, etc.
- Much of the data we use depends on govts and IOs.
- Same issue when states and IOs have low capacity.

Implications for empirical analysis



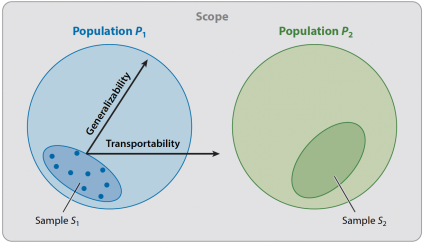
- Treatment $\Rightarrow$ self-selection! (selection into the DV).

Implications for empirical analysis



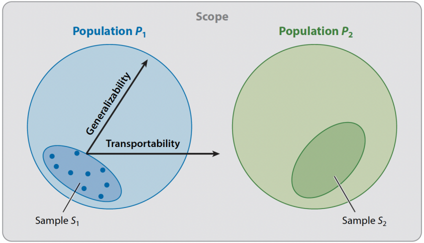
- Treatment $\Rightarrow$ self-selection! (selection into the DV).
- Different from pre-treatment self-selection. Why?

Implications for empirical analysis



- Treatment $\Rightarrow$ self-selection! (selection into the DV).
- Different from pre-treatment self-selection. Why?
- Induces post-treatment self-selection. When? Where?

Implications for empirical analysis



- Treatment $\Rightarrow$ self-selection! (selection into the DV).
- Different from pre-treatment self-selection. Why?
- Induces post-treatment self-selection. When? Where?
- Limits generizability and transportability (scope). Why? When?

Implications for social science

- Institutional change $\Rightarrow$ quality of information.
  - Recall, also induces post-treatment measurement error! Consequences?

Implications for social science

- Institutional change $\Rightarrow$ quality of information.
  - Recall, also induces post-treatment measurement error! Consequences?
- Findings are a function of data quality (limits scope).

Yale

Implications for social science

- Institutional change $\Rightarrow$ quality of information.
  - Recall, also induces post-treatment measurement error! Consequences?
- Findings are a function of data quality (limits scope).
- State, NGO and IO independence can help! But sufficient?
  - Cost of deviation to make it incentive compatible.
  - Signal domestic audience transparency; also for deterrence!
  - It is also a reflection of geopolitics!
    - Differences in resources for collecting quality data.
    - Data manipulation for geopolitical agenda.

Yale

Implications for social science

- Institutional change $\Rightarrow$ quality of information.
  - Recall, also induces post-treatment measurement error! Consequences?
- Findings are a function of data quality (limits scope).
- State, NGO and IO independence can help! But sufficient?
  - Cost of deviation to make it incentive compatible.
  - Signal domestic audience transparency; also for deterrence!
  - It is also a reflection of geopolitics!
    - Differences in resources for collecting quality data.
    - Data manipulation for geopolitical agenda.

Implications for social science

- Institutional change $\Rightarrow$ quality of information.
  - Recall, also induces post-treatment measurement error! Consequences?
- Findings are a function of data quality (limits scope).
- State, NGO and IO independence can help! But sufficient?
  - Cost of deviation to make it incentive compatible.
  - Signal domestic audience transparency; also for deterrence!
  - It is also a reflection of geopolitics!
    - Differences in resources for collecting quality data.
    - Data manipulation for geopolitical agenda.

Imputation and worst case bounds

Table 1. Bias in listwise deletion and multiple imputation.

Missingness	Listwise Deletion	Multiple Imputation
MCAR	Unbiased	Unbiased
MAR (Missing in X)	Unbiased	Unbiased
MAR (Missing in Y, X)	Biased	Unbiased
MNAR/NI (Missing in X)	Unbiased	?
MNAR/NI (Missing in Y, X)	Biased	Biased

- Imputation?

Imputation and worst case bounds

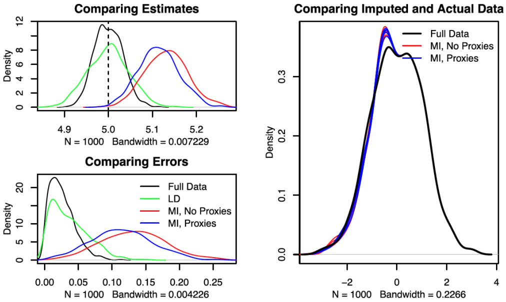
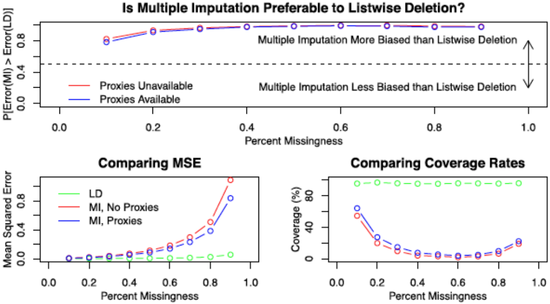


Figure 1. Simulation results.

- Imputation?

Imputation and worst case bounds



- Imputation? Dangerous for addressing selection into DV!

Imputation and worst case bounds

Y_i is our outcome, then

$$R_i = \begin{cases} 1 & \text{Selected,} \\ 0 & \text{otherwise.} \end{cases} \qquad D_i = \begin{cases} 1 & \text{Treated,} \\ 0 & \text{otherwise.} \end{cases} \qquad Y_i = \begin{cases} 1 & \text{Yes,} \\ 0 & \text{No.} \end{cases}$$

- Imputation? Dangerous for addressing selection into DV!
- Manski bounds are a safer alternative; can be uninformative.

Imputation and worst case bounds

Unobserved: $E[Y_{i1}|R_i = 0, D_i = 1]$ and $E[Y_{i0}|R_i = 0, D_i = 0]$

Assume worst:

$$E[Y_{i1}|R_i = 0, Di = 1] = 0 \text{ and } E[Y_{i0}|R_i = 0, Di = 0] = 1$$

$$E[Y_{i1}|R_i = 0, Di = 1] = 1 \text{ and } E[Y_{i0}|R_i = 0, Di = 0] = 0$$

- Imputation? Dangerous for addressing selection into DV!
- Manski bounds are a safer alternative; can be uninformative.

Imputation and worst case bounds

Compute bounds:

$$
\begin{aligned}
BL =& Pr(R_i = 1|Di = 1)E(Y_i|Di = 1, R_i = 1) \\
&- [Pr(R_i = 1|Di = 0)E(Y_i|Di = 0, R_i = 1) + Pr(R_i = 0|Di = 0)]
\end{aligned}
$$

$$
\begin{aligned}
BU =& [Pr(R_i = 1|Di = 1)E(Y_i|Di = 1, R_i = 1) + Pr(R_i = 0|Di = 1)] \\
&- Pr(R_i = 1|Di = 0)E(Y_i|Di = 0, R_i = 1)
\end{aligned}
$$

- Imputation?
- Manski bounds are a safer alternative; can be uninformative.

Imputation and worst case bounds

Trim bounds (never attriters):

$$width = P(R_i = 0|Di = 1) + P(R_i = 0|Di = 0)$$

to

$$width = \frac{Pr(R_i = 1|Di = 1) - Pr(R_i = 1|Di = 0)}{P(R_i = 1|Di = 0)}$$

- Imputation?
- Manski bounds are a safer alternative; can be uninformative.
- Trimming bounds helps but needs (strong) assumptions.
  - Assumption: treatment has an effect on response.
  - If diff. in response rates are small, bounds are informative.

Yale

Next class...

Leaders and bureaucrats!