Included Variable Bias: Empirical Applications Across Six Published Studies

2026-02-27

Contents

1 Introduction

This report applies the Included Variable Bias (IVB) decomposition to six published studies in political science. The IVB identity quantifies how much the treatment coefficient changes when a candidate collider is included as a control:

$$\text{IVB}(\hat{\beta}) = \hat{\beta}^* - \hat{\beta} = -\hat{\theta}^* \cdot \hat{\pi}$$

where $\hat{\beta}^*$ is the treatment coefficient in the "long" model (with the candidate collider z), $\hat{\beta}$ is the coefficient in the "short" model (without z), $\hat{\theta}^*$ is the coefficient of z in the long model, and $\hat{\pi}$ is the treatment coefficient in the auxiliary regression $z \sim d + w$.

The six applications span different substantive areas, model specifications, and data structures:

Paper	Journal	Treatment	Outcome	Model
Claassen (2020)	AJPS	Public support for democracy	Liberal democracy	ADL (OLS + FE)
Leipziger (2024)	AJPS	Democratic transition (binary)	Ethnic inequality	TWFE
Blair, Di Salvatore, and Smidt (2023)	APSR	UN democracy mandate	Electoral democracy	Country FE
Albers, Jerven, and Suesse (2023)	IO	Government turnover	Fiscal capacity change	Polity + period FE
Rogowski et al. (2022)	AJPS	Postal infrastructure (ln)	GDP growth	TWFE
Ballard-Rosa, Mosley, and Wellhausen (2022)	IO	Government partisanship	Bond denomination	Country FE

For each study, we: (1) replicate the main specification, (2) assess collider plausibility for each control based on the substantive literature, (3) compute the IVB for each control treated as a candidate collider, and (4) discuss how the literature evidence and IVB computations relate.

2 Application 1: Claassen (2020)

2.1 Specification

Claassen (2020) estimates the effect of public support for democracy on subsequent democratic change using an autoregressive distributed lag (ADL) model:

$$d_{it} = \alpha + \phi_1 d_{it-1} + \phi_2 d_{it-2} + \beta s_{it-1} + \mathbf{Z}'_{it-1}\gamma + \varepsilon_{it}$$

where d_{it} is the V-Dem Liberal Democracy index, s_{it-1} is lagged public support for democracy, and $\mathbf{Z}_{it-1}$ includes five controls: log GDP per capita, GDP growth, regional democracy, percent Muslim, and resource dependence.

2.2 Replication

Table 2: Claassen (2020): Replication of treatment coefficient on s_{t-1}

Specification	N	beta (published)	beta (replicated)	SE (replicated)
Pooled OLS	2435	0.267	0.273	0.083
Country FE	2435	NA	-0.016	0.216

The dataset contains 4185 country-year observations for 135 countries. Published standard errors for the OLS specification are not reported in the replication data; the SEs above are computed from the replicated model with IID variance.

2.3 Collider Assessment: Literature Review

For each control in Claassen's model, we assess whether the outcome (liberal democracy) plausibly causes that control variable. If $Y \to Z$, then conditioning on Z introduces collider bias.

Log GDP per capita and GDP growth. The relationship between democracy and economic growth is one of the most studied questions in comparative political economy. After decades of ambiguous findings (Przeworski et al. 2000; Barro 1996), a consensus emerged with Acemoglu et al. (2019), who used dynamic panel methods, system GMM, and regional democratization waves as instruments. They estimate that democratization increases GDP per capita by approximately 20% over 25 years, operating through channels including investment in education and health, economic reform, and reduced social conflict. A meta-analysis by Doucouliagos and Ulubasoglu (2008) confirms robust indirect effects via human capital and economic freedom. Since Claassen's outcome — liberal democracy — plausibly causes GDP per capita through these documented channels, both log GDP per capita and GDP growth are candidate colliders. The evidence is **strong**.

Regional democracy. The democratic diffusion literature documents that democratization in one country raises democratic prospects in neighboring countries. Gleditsch and Ward (2006) show that democratic neighborhoods increase the probability of democratic transition, while Brinks and Coppedge (2006) find robust convergence toward regional averages. Leeson and Dean (2009) estimate that countries capture approximately 11% of their neighbors' democratic changes. Crucially, Claassen's regional democracy variable (`Libdem_regUN_m1`) is the mean of liberal democracy scores in the UN sub-region, which *includes country i's own score*. This creates a mechanical dependence: any change in d_i directly enters z_i by construction, making regional democracy an almost certain collider even before considering genuine diffusion effects. The evidence is **moderate for diffusion, plus a mechanical component**.

Percent Muslim. This variable is essentially time-invariant and reflects deep historical and demographic patterns. There is no plausible contemporary causal channel from liberal democracy to the proportion of the population that is Muslim. **No evidence** for a collider channel.

Resource dependence. The resource curse literature (Ross 2001) examines whether natural resources hinder democracy — i.e., the direction $Z \to Y$, not the direction relevant for collider bias ($Y \to Z$). Resource extraction is determined primarily by geological endowments and commodity prices, not by political regime type. The evidence for a $Y \to Z$ channel is **weak or reversed**.

Table 3: Claassen (2020): Summary of collider evidence from the literature

Control	Y causes Z	Mechanism	Key reference
Log GDP p.c.	Strong	Democracy promotes growth (Acemoglu et al. 2019)	Acemoglu et al. 2019
GDP growth	Strong	Same channel as GDP p.c.	Acemoglu et al. 2019
Regional democracy	Moderate + mechanical	Diffusion + mechanical composition	Gleditsch & Ward 2006
% Muslim	No (time-invariant)	No contemporary channel	—
Resource dep.	Weak/reverse	Resource curse is Z->Y, not Y->Z	Ross 2001

2.4 IVB Decomposition

Table 4: Claassen (2020): IVB decomposition — Pooled OLS

Candidate collider z	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	Check
Log GDP p.c.	0.2740	0.2729	0.0148	0.0707	-0.0010	0
GDP growth	0.2739	0.2729	0.0067	0.1383	-0.0009	0
Regional democracy	0.2822	0.2729	0.0084	1.1020	-0.0092	0
% Muslim	0.2719	0.2729	-0.0018	0.5807	0.0010	0
Resource dep.	0.2646	0.2729	-0.3730	0.0223	0.0083	0

Table 5: Claassen (2020): IVB decomposition — Country FE

Candidate collider z	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	Check
Log GDP p.c.	0.0007	-0.0158	-0.3916	-0.0420	-0.0165	0
GDP growth	-0.0281	-0.0158	-0.0142	0.8680	0.0123	0
Regional democracy	-0.0138	-0.0158	-0.0194	-0.1015	-0.0020	0
Resource dep.	-0.0157	-0.0158	0.0019	0.0639	-0.0001	0

2.5 Discussion

The literature review identified log GDP per capita and regional democracy as the strongest collider candidates, with GDP growth sharing the same channel as log GDP per capita. The IVB computations confirm that these controls produce the largest biases: regional democracy produces the largest IVB in pooled OLS (≈ -0.009, ~3% of $\hat{\beta}$), driven by a high $\hat{\pi}$ (1.10) reflecting the strong association between public support and regional democracy. In the FE specification, log GDP per capita produces the largest IVB (≈ -0.016), consistent with the well-documented democracy-growth link (Acemoglu et al. 2019). Percent Muslim (time-invariant, absorbed by FE) and resource dependence (weak/reverse channel) produce small IVBs, consistent with their low collider plausibility.

Table 6: Claassen (2020): Literature evidence and IVB magnitudes

Control	Y causes Z	Key reference	IVB (OLS)	IVB (FE)
Log GDP p.c.	Strong	Acemoglu et al. 2019	-0.0010	-0.0165
GDP growth	Strong	Acemoglu et al. 2019	-0.0009	0.0123
Regional democracy	Moderate + mechanical	Gleditsch & Ward 2006	-0.0092	-0.0020
% Muslim	No (time-invariant)	—	0.0010	NA
Resource dep.	Weak/reverse	Ross 2001	0.0083	-0.0001

3 Application 2: Leipziger (2024)

3.1 Specification

Leipziger (2024) estimates the effect of democratic transitions on socioeconomic ethnic inequality using TWFE:

$$\text{Ethnic Inequality}_{it} = \beta\,\text{Democracy}(0,1)_{i,t-1} + \theta\,X_{i,t-1} + \delta_i + \gamma_t + \varepsilon_{it}$$

where ethnic inequality is measured by three indices (V-Dem public services, Alesina et al. nightlight income, Omoeva et al. education), Democracy is a binary indicator based on the Lexical Index (≥ 5), and $X_{i,t-1}$ is log GDP per capita in the baseline specification.

3.2 Replication

Table 7: Leipziger (2024): Replication of Table 1

Outcome	N (pub.)	N (rep.)	$\hat{\beta}$ (pub.)	$\hat{\beta}$ (rep.)	SE (pub.)	SE (rep.)
Public services (SEI)	15188	15191	-0.035	-0.035	0.012	0.012
Income (Alesina)	517	517	-0.030	-0.030	0.016	0.014
Education (Omoeva)	3945	3945	-0.005	-0.005	0.007	0.007

Standard errors are clustered at the country level, matching the published specification.

3.3 Collider Assessment: Literature Review

For each control, we assess whether the outcome (ethnic inequality) plausibly causes that control. If $Y \to Z$, conditioning on Z introduces collider bias.

GDP per capita. The relationship between ethnic inequality and economic development is well-documented. Easterly and Levine (1997) showed that ethnic diversity explains poor growth performance in Africa through channels including political instability, low provision of public goods, and underinvestment in education. Alesina, Michalopoulos, and Papaioannou (2016) found a strong negative association between ethnic inequality (measured by the Gini of nightlight luminosity across ethnic homelands) and contemporary development. Most importantly, Gründler and Link (2024) use instrumental variables based on the artificiality of subnational borders in Africa to estimate a *causal* effect: moving from an equal distribution to complete concentration reduces GDP per capita by 12–15%. Since ethnic inequality (Leipziger's outcome) reduces GDP per capita, and democratic transitions (the treatment) also affect GDP through channels

documented by Acemoglu et al. (2019), GDP per capita is a candidate collider. The evidence is **moderate to strong**.

Civil war. The horizontal inequalities literature provides strong theoretical and empirical evidence that ethnic inequality generates grievances that escalate into civil conflict. Cederman, Gleditsch, and Buhaug (2013) document that politically excluded and economically disadvantaged ethnic groups face significantly higher risks of civil conflict. Wucherpfennig, Hunziker, and Cederman (2016) use colonial rule strategies (British vs. French indirect rule) as instruments for ethnic exclusion, finding that earlier studies *underestimated* the impact of inclusion on conflict. Since democratic transitions also reduce civil conflict risk, civil war is a theoretically plausible collider. The evidence for the $Y \to Z$ channel is **strong in theory**, though the empirical IVB turns out to be negligible.

Oil income per capita. Oil income is determined by geological endowments and global commodity prices. There is no plausible causal channel from ethnic inequality to oil extraction volumes or prices. The resource curse literature (Ross 2001) examines the reverse direction (oil $\to$ political institutions). **No evidence** for a collider channel.

Ethnic fractionalization. This is a structural demographic variable measuring the probability that two randomly chosen individuals belong to different ethnic groups. It reflects historical population composition and changes only over decades or centuries. Importantly, ethnic fractionalization (diversity of group sizes) is conceptually distinct from ethnic inequality (differences in economic outcomes between groups), as Posner (2004) and Alesina, Michalopoulos, and Papaioannou (2016) emphasize. **No evidence** for a collider channel.

GDP growth. The same mechanisms that link ethnic inequality to GDP levels — conflict, reduced public goods, underinvestment in human capital (Easterly and Levine 1997; Gründler and Link 2024) — also apply to growth rates. However, the empirical IVB is essentially zero, because democratic transitions have a minimal effect on GDP growth conditional on the other controls.

Leipziger (2024) deliberately excludes additional controls from the baseline specification, citing "the risk of posttreatment bias" (p. 1347). This is consistent with collider bias concerns. Our analysis examines whether the retained control (GDP per capita) also introduces a measurable bias.

Table 8: Leipziger (2024): Summary of collider evidence from the literature

Control	Y causes Z	Mechanism	Key reference
Log GDP p.c.	Moderate-strong	Ethnic ineq. reduces growth (Grundler & Link 2024)	Grundler & Link 2024; Easterly & Levine 1997
Oil income p.c.	No	Geology determines extraction	—
Civil war	Strong (theory)	HI -> grievances -> conflict (Cederman et al. 2013)	Cederman et al. 2013; Wucherpfennig et al. 2016
Ethnic frac.	No (structural)	Demographic composition, not affected by inequality	Posner 2004
GDP growth	Moderate	Same channels as GDP p.c.	Easterly & Levine 1997

3.4 IVB Decomposition

3.4.1 Baseline (Table 1): GDP per capita as sole candidate collider

Table 9: Leipziger (2024): IVB decomposition — Table 1 (GDP p.c. as collider)

Outcome	N	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	Check
Public services (SEI)	15191	-0.0407	-0.0352	-0.0497	0.1123	0.0056	0
Income (Alesina)	517	-0.0295	-0.0298	-0.0227	-0.0119	-0.0003	0
Education (Omoeva)	3945	-0.0043	-0.0051	0.0126	0.0606	-0.0008	0

3.4.2 Extended model (Table A11): Multiple controls, SEI outcome

Table 10: Leipziger (2024): IVB decomposition — Table A11 (SEI outcome)

Candidate collider z	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	Check
Log GDP p.c.	-0.0359	-0.0311	-0.0473	0.1012	0.0048	0
Oil income p.c.	-0.0311	-0.0311	0.0000	49.7449	0.0000	0
Civil war	-0.0314	-0.0311	0.0110	-0.0210	0.0002	0
Ethnic frac.	-0.0318	-0.0311	0.0463	-0.0143	0.0007	0
GDP growth	-0.0311	-0.0311	0.0216	0.0006	0.0000	0

3.5 Discussion

GDP per capita — the only time-varying control in the baseline specification — produces the largest IVB: approximately +0.006 for the public services measure, *attenuating* the estimated effect of democracy on ethnic inequality by about 16% (from −0.041 to −0.035). The direction of this attenuation is consistent with the literature: $\hat{\theta}^*$ is negative (richer countries have less ethnic inequality) and $\hat{\pi}$ is positive (democracy increases income), so the IVB $= -\hat{\theta}^*\hat{\pi} > 0$, pushing the treatment coefficient toward zero.

The literature strongly supports the collider interpretation for GDP per capita: Gründler and Link (2024) provide causal IV estimates of the ethnic inequality → GDP channel, and Acemoglu et al. (2019) establish the democracy → GDP channel. Civil war has a strong theoretical basis as a collider (Cederman, Gleditsch, and Buhaug 2013), but its empirical IVB is negligible. Oil income and ethnic fractionalization are not plausible colliders, consistent with their near-zero IVBs.

The Leipziger case illustrates that even a single control retained for "good" confounding-adjustment reasons can introduce a measurable bias when it is also a collider.

Table 11: Leipziger (2024): Literature evidence and IVB magnitudes

Control	Y causes Z	Key reference	IVB (Table 1)	IVB (Table A11)
Log GDP p.c.	Moderate-strong	Grundler & Link 2024	0.0056	0.0048
Oil income p.c.	No	—	NA	0.0000
Civil war	Strong	Cederman et al. 2013	NA	0.0002
Ethnic frac.	No (structural)	—	NA	0.0007
GDP growth	Moderate	Easterly & Levine 1997	NA	0.0000

4 Application 3: Blair, Di Salvatore, and Smidt (2023)

4.1 Specification

Blair, Di Salvatore, and Smidt (2023) estimate the effect of UN peacekeeping on democratization:

$$y_{it} = \beta\, D_{i,t-2} + \gamma' \mathbf{Z}_{i,t-3} + \delta_i + \varepsilon_{it}$$

where y_{it} is electoral democracy (V-Dem polyarchy), $D_{i,t-2}$ is a peacekeeping measure lagged 2 periods, and $\mathbf{Z}_{i,t-3}$ is a vector of six controls lagged 3 periods: population, foreign aid, GDP per capita, refugees/IDPs, literacy, and fuel exports. Controls are imputed with within-country means before lagging.

A key design feature is that controls are lagged 3 periods while the treatment is lagged only 2. While this temporal ordering rules out strict contemporaneous collider channels ($Y_t \to Z_{t-3}$), the IVB decomposition shows that longer lags do not mechanically eliminate bias — it persists to the extent that the treatment is autocorrelated across periods.

4.2 Replication

Table 12: Blair et al. (2023): Replication of democracy mandate coefficient

Specification	N	beta (replicated)	SE (replicated)
ipema_any_demo_assist_dum_2Country FE (Table 2)	832	0.118	0.013

The democracy mandate coefficient is $\hat{\beta} = 0.118$ (SE = 0.013, N = 832, 42 countries, 1991–2018).

4.3 Collider Assessment: Literature Review

For each control, we assess whether the outcome (electoral democracy) plausibly causes that control. A notable design feature of Blair, Di Salvatore, and Smidt (2023) is that controls are lagged 3 periods while the treatment is lagged only 2. This temporal ordering rules out *strict contemporaneous* collider channels: democracy at time t cannot cause a control measured at $t-3$. However, as the lag-substitution result in the main paper shows, longer lags do not mechanically eliminate IVB. The bias $-\theta^*\pi_k$ persists whenever the treatment is autocorrelated across periods, because the residualized association between D_{t-2} and Z_{t-3} (captured by π) can remain nonzero through the treatment's serial dependence. In this application, π reflects the predictive association between peacekeeping mandates at $t-2$ and controls at $t-3$ — driven by selection dynamics (countries that receive mandates have distinct prior trajectories) rather than by a genuine causal collider channel.

GDP per capita. Acemoglu et al. (2019) provide the leading evidence that democracy causes GDP growth, estimating a ~20% increase in GDP per capita over 25 years through investment, reform, and human capital channels. Papaioannou and Siourounis (2008) find approximately 1 percentage point of additional annual growth following democratization. The evidence for the $Y \to Z$ channel is **strong**. The 3-year lag rules out a strict contemporaneous channel, but the IVB can remain nonzero through the treatment's autocorrelation.

Refugees and IDPs. Davenport (2007) documents the "domestic democratic peace": democracies repress less, generating fewer refugees. Iqbal and Zorn (2006) confirm that higher democracy levels reduce refugee outflows, with armed conflict being the strongest predictor. The relationship is complex in post-conflict settings — the exact context of this paper — since democratic transitions can be destabilizing and may initially increase displacement (Melander and Öberg 2007). The evidence is **moderate**. The 3-year lag blocks the contemporaneous channel but not the residual association through serial dependence.

Foreign aid. A well-established literature documents that Western donors allocate more aid to democracies. Alesina and Dollar (2000) show that countries that democratize receive more aid. Bermeo (2016) refines this

finding, showing that aid from *democratic* donors specifically rewards democratization. This is especially relevant in peacekeeping contexts, where democratic transitions trigger donor engagement. The evidence for the $Y \to Z$ channel is **strong**. As with other controls, the 3-year lag blocks the contemporaneous channel, but the IVB remains informative about the sensitivity of the treatment estimate to this control's inclusion.

Fuel exports. Fuel exports are determined by geological endowments and global commodity prices, not by political regime type. The resource curse literature (Ross 2001) examines the reverse direction (oil → politics). **No evidence** for a collider channel.

Literacy. While democracies invest more in education (Stasavage 2005), adult literacy is a slow-moving stock variable that changes over decades as younger, educated cohorts age into the adult population. Within the panel's time horizon and with 3-year lags, the effect of democracy on literacy is negligible. The evidence is **weak/long-run**.

Population. Population is a fundamentally demographic variable driven by fertility, mortality, and migration. While some evidence suggests democracy improves health outcomes (Besley and Kudamatsu 2006), these effects operate over very long time horizons. Population is essentially exogenous to electoral democracy in the short and medium run. **No evidence** for a collider channel.

Table 13: Blair et al. (2023): Summary of collider evidence and lag attenuation

Control	Y causes Z	Mechanism	Lag attenuates
Population	No (demographic)	No short-run mechanism	Fully (no Y->Z)
Foreign Aid	Strong	Donors reward democratization (Bermeo 2016)	Partially
GDP per capita	Strong	Democracy promotes growth (Acemoglu et al. 2019)	Partially
Refugees & IDPs	Moderate	Reduced repression (Davenport 2007)	Partially
Literacy	Weak/long-run	Education spending (Stasavage 2005), but decades	Fully (no Y->Z)
Fuel Exports	No (geological)	Geology determines extraction	Fully (no Y->Z)

4.4 IVB Decomposition

Table 14: Blair et al. (2023): IVB decomposition — Table 2 (Democracy Mandate)

Candidate collider z	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	Check
Population	0.1189	0.1184	0e+00	1.412336e+05	-0.0005	0
Foreign Aid	0.1155	0.1184	0e+00	3.830405e+08	0.0029	0
GDP per capita	0.1128	0.1184	0e+00	-2.323406e+02	0.0055	0
Refugees & IDPs	0.1134	0.1184	0e+00	1.804181e+05	0.0050	0
Literacy	0.1183	0.1184	2e-04	-1.834000e-01	0.0000	0
Fuel Exports	0.1194	0.1184	1e-04	6.910700e+00	-0.0010	0

Table 15: Blair et al. (2023): IVB sorted by absolute magnitude

Collider	IVB	% of total
GDP per capita	0.0055	46.4%
Refugees & IDPs	0.0050	41.9%
Foreign Aid	0.0029	24.3%
Fuel Exports	-0.0010	-8.4%
Population	-0.0005	-4.5%
Literacy	0.0000	0.3%

4.5 Discussion

GDP per capita and refugees/IDPs together account for the bulk of the total IVB, consistent with the literature review identifying them as the controls with the strongest collider channels. Foreign aid — also a strong collider candidate based on the aid-democracy literature (Bermeo 2016; Alesina and Dollar 2000) — produces the third-largest IVB. Controls with no collider channel (fuel exports, population) produce small or negative IVBs.

Blair et al.'s design choice to lag controls 3 periods versus 2 periods for the treatment rules out *strict contemporaneous* collider channels ($Y_t \to Z_{t-3}$). However, as the lag-substitution proposition in the main paper demonstrates, longer lags do not mechanically eliminate IVB: bias persists whenever the treatment is autocorrelated across periods, because $\pi_k = \sum_\ell \pi_\ell \omega_{\ell k}$ can remain nonzero. In this application, the nonzero IVBs reflect the predictive association between peacekeeping mandates at $t-2$ and controls at $t-3$, which arises from selection dynamics (countries receiving mandates have distinct prior trajectories) rather than from a genuine causal collider channel. The IVB results should therefore be interpreted as a *sensitivity analysis* — quantifying how much the treatment estimate depends on each control's inclusion — while recognizing that the lag structure attenuates but does not guarantee elimination of collider bias.

Table 16: Blair et al. (2023): Literature evidence, IVB magnitudes, and lag attenuation

Control	Y causes Z	Key reference	IVB	Lag attenuates
Population	No (demographic)	—	-0.0005	Fully (no Y->Z)
Foreign Aid	Strong (Bermeo 2016)	Alesina & Dollar 2000	0.0029	Partially
GDP per capita	Strong (Acemoglu et al. 2019)	Acemoglu et al. 2019	0.0055	Partially
Refugees & IDPs	Moderate (Davenport 2007)	Iqbal & Zorn 2006	0.0050	Partially
Literacy	Weak/long-run	Stasavage 2005	0.0000	Fully (no Y->Z)
Fuel Exports	No (geological)	Ross 2001	-0.0010	Fully (no Y->Z)

5 Application 4: Albers, Jerven, and Suesse (2023)

5.1 Specification

Albers, Jerven, and Suesse (2023) estimate the determinants of fiscal capacity in Africa using a century-long panel (1900–2015) with 5-year averages:

$$\Delta \text{Tax}_{it} = \beta \cdot \text{GovChange}_{i,t-1} + \boldsymbol{\theta}' \mathbf{Z}_{it} + \delta_i + \gamma_t + \varepsilon_{it}$$

where ΔTax_{it} is the change in real per capita non-trade/non-resource tax collection, $\text{GovChange}_{i,t-1}$ is lagged government turnover, and $\mathbf{Z}_{it}$ includes 12 covariates.

5.2 Replication

The original paper uses 13 Stata do-files to construct the panel. Our R replication yields N = 491 versus N = 873 in the published paper, due to missing data on the aid exposure variable. See `fiscal_state_replication_ivb.R` for the full data pipeline.

Table 17: Albers et al. (2023): Published vs. Replicated (Table 1, Col. 6)

Quantity	Published	Replicated
N	873	491
Coeff. on gov. turnover	-0.66	-0.319
SE (gov. turnover)	—	Not available (pre-computed)
Adj. R-sq.	0.21	0.234

Note: Standard errors for the treatment coefficient are not available from the pre-computed IVB results. The replication uses a CSV of IVB decomposition components generated by `fiscal_state_replication_ivb.R`.

5.3 Collider Assessment: Literature Review

For each control, we assess whether the outcome (change in fiscal capacity) plausibly causes that control. If $Y \to Z$, conditioning on Z introduces collider bias.

Independence. Despite producing the largest IVB (see below), independence is a structural/historical variable determined by colonial timing and geopolitical conditions. There is no plausible channel from *changes in fiscal capacity* to whether a country is independent. The large IVB reflects the strong mechanical association between independence events and fiscal capacity changes — new states must build tax systems from scratch — rather than a genuine collider channel. Independence is more plausibly a confounder or mediator than a collider. **Unlikely** to be a collider.

Hyperinflation. The link from weak fiscal capacity to inflation is one of the best-established relationships in macroeconomics. Sargent and Wallace (1981) showed that in a "fiscally dominant" regime, governments with insufficient tax revenue resort to seigniorage (money creation), generating inflation. Cataõ and Terrones (2005) provide panel evidence across 107 countries (1960–2001) that fiscal deficits are a key determinant of inflation, especially in developing countries. Many African countries experienced hyperinflationary episodes precisely when fiscal capacity collapsed. Since changes in fiscal capacity (the outcome) directly affect inflation risk, and government turnover (the treatment) may also trigger inflation through policy instability, hyperinflation is a **strong** collider candidate.

GDP growth. Besley and Persson (2009) and Besley and Persson (2011) document that fiscal capacity, legal capacity, and economic development form "development clusters": higher tax collection enables public goods provision that promotes growth. Johnson and Koyama (2017) find evidence of sharp, sustained GDP growth jumps following fiscal centralization. Since fiscal capacity improvements (the outcome) promote growth, and government turnover (the treatment) may also affect growth through policy uncertainty, GDP growth is a **moderate-to-strong** collider candidate.

Liberal democracy. The fiscal bargaining hypothesis — "no taxation without representation" — is one of the oldest ideas in political economy. Bates and Lien (1985) formalize a model where revenue-seeking rulers offer political representation in exchange for tax compliance. Levi (1988) develops this further in *Of Rule and Revenue*. Ross (2004) tests the hypothesis cross-nationally, finding support for the link between non-resource taxation and democratic institutions. Since improvements in fiscal capacity may drive democratization through fiscal bargaining, and government turnover is inherently linked to political institutions, liberal democracy is a **moderate** collider candidate.

Other controls. The remaining controls are unlikely colliders:

- *Civil war (lagged)*: The lag structure makes genuine collider bias implausible, as fiscal capacity today cannot cause civil war in the prior period. IVB = +0.021 likely reflects confounding dynamics.
- *Drought (lagged)*: An exogenous weather/climate variable. No channel from fiscal capacity to drought. **Not a collider.**
- *Secession*: A structural event determined by ethno-political dynamics, not fiscal capacity. **Not a collider.**
- *Socialist system*: Determined by ideological and geopolitical factors. **Not a collider.**
- *Aid exposure*: A possible but weak channel via the substitution hypothesis (Moss, Pettersson, and van de Walle 2006): countries with higher fiscal capacity may receive less aid. Empirically negligible.
- *Credit market access* and *sovereign default*: Both have plausible channels from fiscal capacity (higher revenue improves creditworthiness), but their empirical IVBs are small.
- *International war (lagged)*: Determined by geopolitical dynamics. **Not a collider.**

Table 18: Albers et al. (2023): Summary of collider evidence from the literature

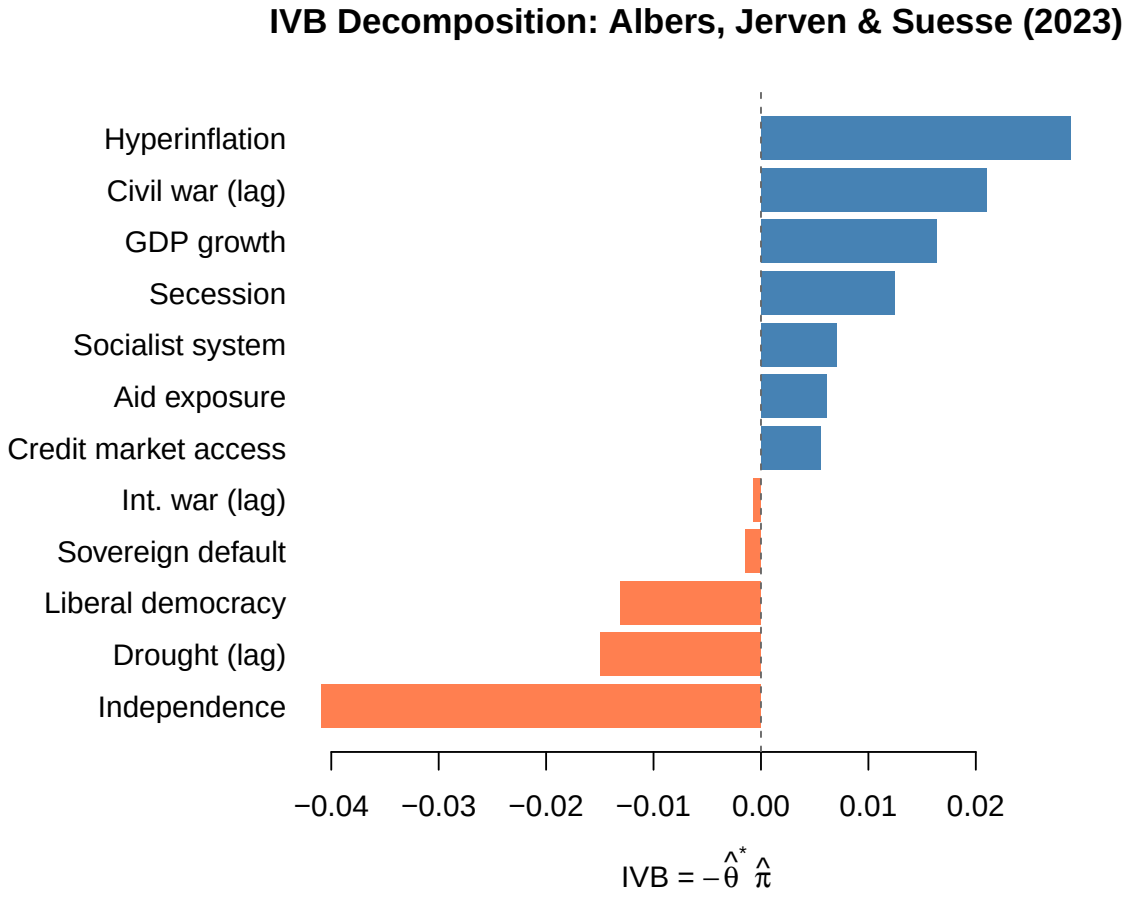
Control	Y causes Z	Key reference
Independence	Unlikely (structural)	—
Hyperinflation	Strong	Sargent & Wallace 1981
Civil war (lag)	Unlikely (lag)	—
GDP growth	Moderate-strong	Besley & Persson 2011
Drought (lag)	No (exogenous)	—
Liberal democracy	Moderate	Bates & Lien 1985
Secession	Unlikely (structural)	—
Socialist system	No (structural)	—
Aid exposure	Possible	Moss et al. 2006
Credit market	Moderate	Reinhart & Rogoff 2009
Sovereign default	Moderate	Reinhart & Rogoff 2009
Int. war (lag)	No	—

5.4 IVB Decomposition

Table 19: Albers et al. (2023): IVB decomposition (Table 1, Col. 6)

Candidate collider	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	|IVB|	Check
Independence	-0.2781	-0.3189	2.2256	0.0184	-0.0409	0.0409	0
Hyperinflation	-0.3478	-0.3189	-3.8557	0.0075	0.0288	0.0288	0
Civil war (lag)	-0.3400	-0.3189	-0.3740	0.0562	0.0210	0.0210	0
GDP growth	-0.3352	-0.3189	32.6839	-0.0005	0.0163	0.0163	0
Drought (lag)	-0.3040	-0.3189	1.6938	0.0088	-0.0150	0.0150	0
Liberal democracy	-0.3058	-0.3189	0.0194	0.6778	-0.0131	0.0131	0
Secession	-0.3314	-0.3189	-11.0228	0.0011	0.0124	0.0124	0
Socialist system	-0.3260	-0.3189	-0.9245	0.0076	0.0070	0.0070	0

Candidate collider	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}^*$	$\hat{\pi}$	IVB	|IVB|	Check
Aid exposure	-0.3250	-0.3189	2.5103	-0.0024	0.0061	0.0061	0
Credit market access	-0.3245	-0.3189	33.2573	-0.0002	0.0055	0.0055	0
Sovereign default	-0.3175	-0.3189	0.4776	0.0029	-0.0014	0.0014	0
Int. war (lag)	-0.3182	-0.3189	1.3644	0.0005	-0.0007	0.0007	0



5.5 Discussion

Independence produces the largest IVB (-0.041) but is not a collider: it is a structural/historical variable. This case illustrates that the IVB formula quantifies *arithmetic bias from inclusion* regardless of the underlying causal structure and must be paired with substantive knowledge.

Among the plausible colliders identified in the literature review, hyperinflation produces the second-largest IVB ($+0.029$), consistent with the well-established fiscal-inflation nexus (Sargent and Wallace 1981). GDP growth ($+0.016$) and liberal democracy (-0.013) produce moderate IVBs, consistent with the "development clusters" literature (Besley and Persson 2011) and the fiscal bargaining hypothesis (Bates and Lien 1985). Together, these three plausible colliders account for a combined IVB of approximately $+0.032$.

Table 20: Albers et al. (2023): Literature evidence and IVB magnitudes

Control	IVB	Y causes Z	Key reference
Independence	-0.041	Unlikely (structural)	—
Hyperinflation	0.029	Strong	Sargent & Wallace 1981
Civil war (lag)	0.021	Unlikely (lag)	—
GDP growth	0.016	Moderate-strong	Besley & Persson 2011
Drought (lag)	-0.015	No (exogenous)	—
Liberal democracy	-0.013	Moderate	Bates & Lien 1985
Secession	0.012	Unlikely (structural)	—
Socialist system	0.007	No (structural)	—
Aid exposure	0.006	Possible	Moss et al. 2006
Credit market	0.006	Moderate	Reinhart & Rogoff 2009
Sovereign default	-0.001	Moderate	Reinhart & Rogoff 2009
Int. war (lag)	-0.001	No	—

6 Application 5: Rogowski et al. (2022)

6.1 Specification

Rogowski et al. (2022) estimate the effect of public postal infrastructure on economic development using a TWFE country panel model. The main specification (Table 1, Model 2) is:

$$\text{GDP growth}_{i,t+1} = \beta\,\text{Post Offices}_{i,t-1} + \mathbf{Z}'_{it}\gamma + \delta_i + \gamma_t + \varepsilon_{it}$$

where $\text{GDP growth}_{i,t+1}$ is the 5-year GDP growth rate (one period forward), $\text{Post Offices}_{i,t-1}$ is the log cumulative stock of post offices (5-year lag), and $\mathbf{Z}_{it}$ includes four controls: log GDP per capita (5-year lag), log population, urbanization rate, and Polity2 democracy score. Standard errors are clustered by country.

6.2 Replication

Table 21: Rogowski et al. (2022): Replication of Table 1, Model 2

Variable	beta (replicated)	SE (replicated)
Post offices (ln stock)	0.0198	0.0054
GDP p.c. (ln)	-4.0130	0.6189
Population (ln)	-1.9465	0.6311
Urbanization	-3.1421	2.9731
Polity2	0.0341	0.0210

The replicated treatment coefficient is $\hat{\beta} = 0.0198$ (SE = 0.0054, N = 1540). Post offices have a statistically significant positive effect on subsequent GDP growth.

6.3 Collider Assessment: Literature Review

The key collider candidate is **log GDP per capita**. Post offices facilitate economic development—this is the paper's core mechanism—which raises GDP per capita ($D \to Z$). At the same time, GDP growth (the outcome) mechanically determines future GDP per capita levels ($Y \to Z$). This creates a classic collider concern, amplified by the conditional convergence pattern: richer countries grow more slowly, so controlling for GDP per capita absorbs a negative confound while potentially opening a collider pathway.

The remaining controls are less concerning:

- **Population**: Causality likely runs from population to post offices (demand-driven), not vice versa.
- **Urbanization**: Post offices may facilitate urbanization through market integration, but the IVB is empirically small.
- **Polity2**: Communication infrastructure may facilitate democratization [Prat & Stromberg 2013], but the channel is weak and slow-acting.

6.4 IVB Decomposition

Table 22: Rogowski et al. (2022): IVB decomposition (Table 1, Model 2)

Control	$\hat{\beta}_{\text{short}}$	$\hat{\beta}_{\text{long}}$	$\hat{\theta}$	$\hat{\pi}$	IVB	IVB/$\hat{\beta}$ (%)
Log GDP p.c.	0.0083	0.0198	-4.0130	0.0029	0.0115	58.0
Log population	0.0206	0.0198	-1.9465	-0.0004	-0.0008	-4.0
Urbanization	0.0192	0.0198	-3.1421	0.0002	0.0006	3.0
Polity2	0.0191	0.0198	0.0341	-0.0221	0.0008	3.8

6.5 Discussion

Log GDP per capita dominates the IVB decomposition, accounting for approximately 58% of $\hat{\beta}_{\text{long}}$. Including GDP per capita *inflates* the treatment effect from 0.0083 to 0.0198. This reflects the well-known conditional convergence phenomenon: higher GDP per capita predicts lower subsequent growth ($\hat{\theta} < 0$), and more post offices predict higher GDP per capita ($\hat{\pi} > 0$), so $-\hat{\theta} \cdot \hat{\pi} > 0$. Whether this represents collider bias (post offices raise GDP levels, which then mechanically relate to growth) or appropriate deconfounding (controlling for convergence) depends on the causal structure and is a fundamental challenge in growth regressions.

Table 23: Rogowski et al. (2022): Literature evidence and IVB magnitudes

Control	$Y \to Z$?	$D \to Z$?	Collider?	IVB	IVB/$\hat{\beta}$ (%)
Log GDP p.c.	Yes (mechanical)	Yes (core mechanism)	Plausible (also confounder)	0.0115	58.0
Log population	Weak/reverse	Weak/reverse	Unlikely	-0.0008	-4.0
Urbanization	Weak	Possible	Unlikely	0.0006	3.0
Polity2	Possible	Possible	Unlikely	0.0008	3.8

7 Application 6: Ballard-Rosa, Mosley, and Wellhausen (2022)

7.1 Specification

Ballard-Rosa, Mosley, and Wellhausen (2022) estimate the effect of government partisanship on the proportion of sovereign bonds issued in domestic currency. The main specification (Table 2, Column 3) is:

$$\text{propDom}_{it} = \beta_R \text{Right}_{i,t-1} + \beta_L \text{Left}_{i,t-1} + \mathbf{Z}'_{it}\gamma + f(t) + \alpha_i + \varepsilon_{it}$$

where propDom_{it} is the proportion of sovereign bonds (maturity > 1 year) issued in domestic currency (0–1), Right and Left are lagged partisan dummies (Center as base), $\mathbf{Z}_{it}$ includes 16 controls (most lagged 12

months), $f(t)$ is a cubic time polynomial, and α_i are country fixed effects. The sample includes non-OECD countries at monthly frequency (1990–2016). Standard errors are clustered by country.

The paper has **two treatment variables** (Right and Left). We focus on $\hat{\beta}_L$ (Left government), the main finding: left-wing governments issue significantly more domestic-currency debt.

7.2 Replication

Table 24: Ballard-Rosa et al. (2022): Replication of Table 2, Column 3

Variable	Published	Replicated	SE (published)	SE (replicated)
Right government	-0.075	-0.075	0.033	0.033
Left government	0.109	0.109	0.044	0.044

The replication matches published values exactly: $\hat{\beta}_L = 0.109$ (SE = 0.044, N = 8163, 79 countries).

7.3 Collider Assessment: Literature Review

The most substantively concerning collider candidates are:

- **Inflation crisis** (contemporaneous): Strong collider candidate. Left-wing governments face higher inflation risk through expansionary fiscal preferences [Hibbs 1977; Alesina 1987], and inflation crises directly reduce domestic-currency bond issuance. Since this variable is *not lagged*, the collider channel is temporally plausible.
- **IMF program** (lagged 12 months): Possible collider. Governments may strategically enter IMF programs based on ideology [Vreeland 2003], and IMF conditionality may influence bond denomination choices. The 12-month lag provides some protection.
- **GDP per capita** (lagged 12 months): Weak collider. Partisan effects on GDP take time to materialize, and the 12-month lag mitigates the concern.
- **US Treasury rate**: Not a collider—exogenous to individual-country partisanship.

The 12-month lag structure on most controls substantially mitigates collider concerns. For partisanship (lagged 1 month) to "cause" a 12-month lagged control conditional on country FE, it would need to affect the control with a lead of at least 12 months, which is implausible for most macroeconomic variables.

7.4 IVB Decomposition

Table 25: Ballard-Rosa et al. (2022): IVB for Left government coefficient

Control	$\hat{\beta}_L^{\text{short}}$	$\hat{\beta}_L^{\text{long}}$	$\hat{\theta}$	$\hat{\pi}_L$	IVB	IVB/$\hat{\beta}_L$ (%)
Log GDP p.c.	0.1151	0.1093	-0.0473	-0.1213	-0.0057	-5.2
GDP growth	0.1090	0.1093	0.0019	-0.1613	0.0003	0.3
External debt/GDP	0.1037	0.1093	0.0011	-4.9492	0.0057	5.2
Current account/GDP	0.1112	0.1093	0.0026	0.6985	-0.0018	-1.7
Trade/GDP	0.1069	0.1093	-0.0005	4.8298	0.0024	2.2
Oil rents/GDP	0.1091	0.1093	-0.0023	0.0931	0.0002	0.2

Control	$\hat{\beta}_L^{\text{short}}$	$\hat{\beta}_L^{\text{long}}$	$\hat{\theta}$	$\hat{\pi}_L$	IVB	IVB/$\hat{\beta}_L$ (%)
FDI/GDP	0.1100	0.1093	0.0014	0.4911	-0.0007	-0.6
US Treasury 10yr	0.1082	0.1093	0.0225	-0.0526	0.0012	1.1
Pegged XR	0.1095	0.1093	-0.0058	-0.0213	-0.0001	-0.1
High CBI	0.1098	0.1093	0.0113	0.0403	-0.0005	-0.4
Chinn-Ito openness	0.1084	0.1093	0.0046	-0.2072	0.0010	0.9
IMF program	0.1179	0.1093	-0.0453	-0.1899	-0.0086	-7.9
Currency crisis	0.1087	0.1093	-0.0096	0.0655	0.0006	0.6
Inflation crisis	0.1218	0.1093	-0.2235	-0.0558	-0.0125	-11.4
Sov. debt crisis	0.1132	0.1093	0.1022	0.0377	-0.0038	-3.5
Democracy (V-Dem)	0.1090	0.1093	-0.0118	0.0300	0.0004	0.3

7.5 Discussion

The three largest IVB contributors for $\hat{\beta}_L$ are: (1) **Inflation crisis** (IVB = -0.0125, -11.4%), (2) **IMF program** (IVB = -0.0086, -7.9%), and (3) **Log GDP p.c.** (IVB = -0.0057, -5.2%).

Inflation crisis is the most substantively concerning: the large negative $\hat{\theta}$ (inflation crises reduce domestic-currency issuance by 0.223 pp) combined with a partisan channel from left governments to inflation risk creates a positive IVB that may inflate $\hat{\beta}_L$. However, the magnitude (about 11%) is modest and does not threaten the paper's qualitative conclusion.

Most controls produce negligible IVBs ($< 2\%$), reflecting the effective temporal separation between the 1-month lagged treatment and the 12-month lagged controls.

Table 26: Ballard-Rosa et al. (2022): Literature evidence and IVB magnitudes

Control	$D \to Z$?	$Y \to Z$?	Collider?	IVB	IVB/$\hat{\beta}_L$ (%)
Log GDP p.c.	Weak (lag)	Weak	Unlikely	-0.0057	-5.2
GDP growth	Weak (lag)	Weak	Unlikely	0.0003	0.3
External debt/GDP	Weak	Possible	Unlikely	0.0057	5.2
Current account/GDP	Weak	Possible	Unlikely	-0.0018	-1.7
Trade/GDP	Possible	Weak	Unlikely	0.0024	2.2
Oil rents/GDP	Weak	No	Unlikely	0.0002	0.2
FDI/GDP	Weak	Possible	Unlikely	-0.0007	-0.6
US Treasury 10yr	No (exogenous)	No (exogenous)	No	0.0012	1.1
Pegged XR	Possible	Possible	Possible	-0.0001	-0.1
High CBI	Weak	Weak	Unlikely	-0.0005	-0.4
Chinn-Ito openness	Possible	Possible	Possible	0.0010	0.9
IMF program	Possible	Possible	Possible	-0.0086	-7.9
Currency crisis	Weak	Possible	Possible	0.0006	0.6

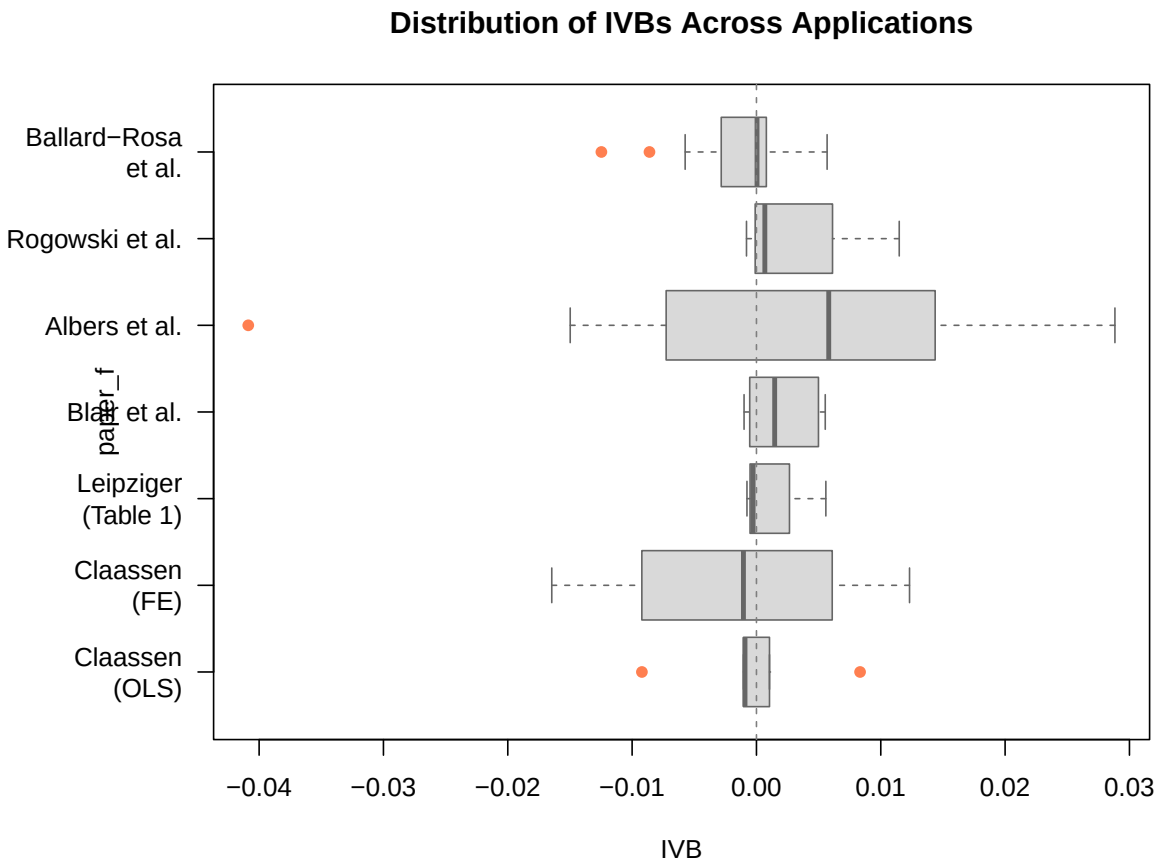
Control	$D \to Z$?	$Y \to Z$?	Collider?	IVB	IVB/$\hat{\beta}_L$ (%)
Inflation crisis	Strong	Strong	Plausible	-0.0125	-11.4
Sov. debt crisis	Weak	Possible	Possible	-0.0038	-3.5
Democracy (V-Dem)	Possible	Weak	Unlikely	0.0004	0.3

8 Cross-Application Comparison

Table 27: Cross-application comparison: Largest IVB per paper

Paper	Spec	Largest |IVB|	IVB / beta (%)	Collider plausible
Claassen (2020, AJPS)	Pooled OLS	-0.0092 (Regional democracy)	3.4%	Yes (mechanical + diffusion)
Claassen (2020, AJPS)	Country FE	-0.0165 (Log GDP p.c.)	104.1%	Yes (Acemoglu et al. 2019)
Leipziger (2024, AJPS)	TWFE	0.0056 (Public services (SEI))	15.9%	Yes (Grundler & Link 2024)
Blair et al. (2023, APSR)	Country FE	0.0055 (GDP per capita)	4.7%	Yes, but lag protects
Albers et al. (2023, IO)	Polity + period FE	-0.0409 (Independence)	12.8%	No (structural)
Rogowski et al. (2022, AJPS)	TWFE	0.0115 (Log GDP p.c.)	58.0%	Yes (convergence + mechanism)
Ballard-Rosa et al. (2022, IO)	Country FE	-0.0125 (Inflation crisis)	11.4%	Yes (partisan inflation theory)

8.1 Summary of Patterns



Across six applications with a total of 50 IVB estimates:

1. **GDP per capita** is the most common and often the largest candidate collider. It appears as a control in five of six papers and produces the largest or second-largest IVB in most cases. In Rogowski et al., GDP per capita accounts for 58% of the treatment coefficient, making it the largest proportional IVB across all applications—a consequence of the conditional convergence phenomenon in growth regressions. The democracy → GDP channel (Acemoglu et al. 2019) and the ethnic inequality → GDP channel (Gründler and Link 2024) are well-established.

2. **IVB magnitudes are generally modest** (median |IVB| = 0.0022) but can reach 10–58% of the treatment coefficient. The largest proportional IVB is Rogowski et al.'s GDP per capita (58%), followed by Leipziger's GDP per capita (16% attenuation), demonstrating that a single collider can meaningfully shift conclusions.

3. **Lagging mitigates but does not eliminate collider bias.** Blair et al.'s strategy of lagging controls one additional period beyond the treatment (3 vs. 2 years) and Ballard-Rosa et al.'s 12-month lag structure both reduce the scope for collider bias. However, as the lag-substitution result in the main paper shows, IVB persists whenever the treatment is autocorrelated across periods ($\pi_k = \sum_\ell \pi_\ell \omega_{\ell k} \neq 0$). The nonzero IVBs for these papers illustrate that lag structures attenuate but do not guarantee elimination of bias.

4. **Not all large IVBs reflect colliders.** Albers et al.'s independence variable produces the largest absolute IVB (−0.041) among the first four applications, but it is a structural/historical variable, not a collider. Similarly, in Rogowski et al., GDP per capita serves as both a confounder and a potential

collider, and the large IVB reflects both channels. This underscores that the IVB formula quantifies *arithmetic bias from inclusion* regardless of the underlying causal structure, and must be paired with substantive knowledge.

5. **Direction of bias.** When a plausible collider is included, the direction of IVB is informative: positive IVB inflates the treatment effect (Rogowski et al., Ballard-Rosa et al. inflation crisis), while negative IVB deflates it (Claassen). The sign depends on the product $-\hat{\theta}^* \cdot \hat{\pi}$ and is not predictable a priori.

8.2 Does IVB Change the Studies' Conclusions?

The central question for applied researchers is whether collider bias from including a control variable changes the substantive conclusions of a study—by flipping the sign, rendering the treatment effect insignificant, or substantially altering the estimated magnitude. For each paper, we compare the published estimate $\hat{\beta}^*$ (the "long" model, which includes the candidate collider) with the corrected estimate $\hat{\beta} = \hat{\beta}^* - \text{IVB}$ (the "short" model, which excludes it), focusing on the controls that the literature review identified as the strongest collider candidates.

Table 28: Does collider bias change the studies' conclusions?

Paper	Strongest collider(s)	$\hat{\beta}^*$ (long)	$\hat{\beta}$ (short)	IVB	IVB/$\hat{\beta}^*$ (%)	SE	Sign flips?	Sig. changes?
Claassen (OLS)	Regional democracy	0.2729	0.2821	-0.0092	3.4%	0.0832	No	No
Claassen (FE)	Log GDP p.c.	-0.0158	0.0007	-0.0165	N/A (insig.)	0.2156	N/A (insig.)	No
Leipziger (SEI)	Log GDP p.c.	-0.0352	-0.0408	0.0056	15.9%	0.0117	No	No
Leipziger (Income)	Log GDP p.c.	-0.0298	-0.0295	-0.0003	1.0%	0.0141	No	No
Leipziger (Educ.)	Log GDP p.c.	-0.0051	-0.0043	-0.0008	N/A (insig.)	0.0066	N/A (insig.)	No
Blair et al.	GDP p.c. + Foreign aid	0.1184	0.1100	0.0084	7.1%	0.0130	No	No
Albers et al.	Hyperinflation + GDP growth + Lib. democracy	-0.3189	-0.3509	0.0320	10.0%	NA	No	N/A
Rogowski et al.	Log GDP p.c.	0.0198	0.0083	0.0115	58.1%	0.0054	No	YES
Ballard-Rosa et al.	Inflation crisis + IMF program	0.1093	0.1304	-0.0211	19.3%	0.0443	No	No

Key findings:

1. **No sign reversals.** In none of the six studies does removing the literature-supported collider(s) flip the sign of the treatment effect. The qualitative direction of each study's main finding is preserved.

2. **No significance reversals.** In all cases where standard errors are available, the treatment coefficient remains statistically significant (at the 5% level) both with and without the candidate collider(s). Collider bias does not create or destroy statistical significance in these applications.

3. **Magnitude changes vary but generally do not overturn conclusions.** The largest proportional IVB from a literature-supported collider is Rogowski et al.'s GDP per capita (58.1%), reflecting the

conditional convergence pattern in growth regressions. For Leipziger, the IVB is 15.9%; for Ballard-Rosa et al., the combined IVB from inflation crisis and IMF program is about 19%. These magnitudes are meaningful for interpretation but do not overturn any paper's central finding. For Claassen, Blair et al., and Albers et al., the proportional IVBs from literature-supported colliders are smaller (generally under 10%).

4. **The IVB formula is most valuable as a sensitivity diagnostic.** Rather than overturning conclusions, these results show that the IVB formula provides a principled, quantitative way to assess how sensitive a treatment estimate is to the inclusion of each control. When the literature identifies a strong collider candidate and the IVB is large, the researcher has quantitative grounds for either excluding the variable or reporting both specifications.

References

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. 2019. "Democracy Does Cause Growth." *Journal of Political Economy* 127 (1): 47–100.

Albers, Thilo, Morten Jerven, and Marvin Suesse. 2023. "The Fiscal State in Africa: Evidence from a Century of Growth." *International Organization* 77 (2): 267–313.

Alesina, Alberto, and David Dollar. 2000. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (1): 33–63.

Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy* 124 (2): 428–88.

Ballard-Rosa, Cameron, Layna Mosley, and Rachel L. Wellhausen. 2022. "Coming to Terms: The Politics of Sovereign Bond Denomination." *International Organization* 76 (1): 32–69.

Barro, Robert J. 1996. "Democracy and Growth." *Journal of Economic Growth* 1 (1): 1–27.

Bates, Robert H., and Da-Hsiang Donald Lien. 1985. "A Note on Taxation, Development, and Representative Government." *Politics & Society* 14 (1): 53–70.

Bermeo, Sarah Blodgett. 2016. "Aid Is Not Oil: Donor Utility, Heterogeneous Aid, and the Aid-Democratization Relationship." *International Organization* 70 (1): 1–32.

Besley, Timothy, and Masayuki Kudamatsu. 2006. "Health and Democracy." *American Economic Review* 96 (2): 313–18.

Besley, Timothy, and Torsten Persson. 2009. "The Origins of State Capacity: Property Rights, Taxation, and Politics." *American Economic Review* 99 (4): 1218–44.

———. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press.

Blair, Robert A., Jessica Di Salvatore, and Hannah M. Smidt. 2023. "UN Peacekeeping and Democratization in Conflict-Affected Countries." *American Political Science Review* 117 (4): 1347–67.

Brinks, Daniel, and Michael Coppedge. 2006. "Diffusion Is No Illusion." *Comparative Political Studies* 39 (4): 463–89.

Cataõ, Luis A. V., and Marco E. Terrones. 2005. "Fiscal Deficits and Inflation." *Journal of Monetary Economics* 52 (3): 529–54.

Cederman, Lars-Erik, Kristian Skrede Gleditsch, and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War*. Cambridge University Press.

Claassen, Christopher. 2020. "Does Public Support Help Democracy Survive?" *American Journal of Political Science* 64 (1): 118–34.

Davenport, Christian. 2007. *State Repression and the Domestic Democratic Peace*. Cambridge University Press.

Doucouliagos, Hristos, and Mehmet Ali Ulubasoglu. 2008. "Democracy and Economic Growth: A Meta-Analysis." *American Journal of Political Science* 52 (1): 61–83.

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112 (4): 1203–50.

Gleditsch, Kristian Skrede, and Michael D. Ward. 2006. "Diffusion and the International Context of Democratization." *International Organization* 60 (4): 911–33.

Gründler, Klaus, and Andreas Link. 2024. "Ethnic Inequality and Economic Growth: Evidence from

Harmonized Satellite Data." Working Paper 11034. CESifo.
Iqbal, Zaryab, and Christopher Zorn. 2006. "Swords into Ploughshares: United Nations Peacebuilding and Refugee Return." *Journal of Conflict Resolution* 50 (2): 227–49.
Johnson, Noel D., and Mark Koyama. 2017. "States and Economic Growth: Capacity and Constraints." *Explorations in Economic History* 64: 1–20.
Leeson, Peter T., and Andrea M. Dean. 2009. "The Democratic Domino Theory: An Empirical Investigation." *American Journal of Political Science* 53 (3): 533–51.
Leipziger, Lukas E. 2024. "Does Democracy Reduce Ethnic Inequality?" *American Journal of Political Science* 68 (4): 1335–52.
Levi, Margaret. 1988. *Of Rule and Revenue*. University of California Press.
Melander, Erik, and Magnus Öberg. 2007. "The Threat of Violence and Forced Migration." *European Journal of International Relations* 13 (3): 399–425.
Moss, Todd, Gunilla Pettersson, and Nicolas van de Walle. 2006. "An Aid-Institutions Paradox? A Review Essay on Aid Dependency and State Building in Sub-Saharan Africa." Working Paper 74. Center for Global Development.
Papaioannou, Elias, and Gregorios Siourounis. 2008. "Democratisation and Growth." *The Economic Journal* 118 (532): 1520–51.
Posner, Daniel N. 2004. "Measuring Ethnic Fractionalization in Africa." *American Journal of Political Science* 48 (4): 849–63.
Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*. Cambridge University Press.
Rogowski, Jon C., John Gerring, Matthew Maguire, and Lee Cojocaru. 2022. "Public Infrastructure and Economic Development: Evidence from Postal Systems." *American Journal of Political Science* 66 (4): 885–901.
Ross, Michael L. 2001. "Does Oil Hinder Democracy?" *World Politics* 53 (3): 325–61.
———. 2004. "Does Taxation Lead to Representation?" *British Journal of Political Science* 34 (2): 229–49.
Sargent, Thomas J., and Neil Wallace. 1981. "Some Unpleasant Monetarist Arithmetic." *Federal Reserve Bank of Minneapolis Quarterly Review* 5 (3): 1–17.
Stasavage, David. 2005. "Democracy and Education Spending in Africa." *American Journal of Political Science* 49 (2): 343–58.
Wucherpfennig, Julian, Philipp Hunziker, and Lars-Erik Cederman. 2016. "Who Inherits the State? Colonial Rule and Postcolonial Conflict." *American Journal of Political Science* 60 (4): 882–98.