RedNIE Red Nacional de Investigadores en Economía

# When Is the Sovereign Greenium a Scarcity Price? Green Supply, Market Depth, and Cross-Country Evidence

**Juan Pablo Brichetti** (IDB)

**Eduardo Cavallo**(IDB)

**Francisco Roch** (UTDT)

DOCUMENTO DE TRABAJO N° 413

Noviembre de 2026

Los documentos de trabajo de la RedNIE se difunden con el propósito de generar comentarios y debate, no habiendo estado sujetos a revisión de pares. Las opiniones expresadas en este trabajo son de los autores y no necesariamente representan las opiniones de la RedNIE o su Comisión Directiva.

The RedNIE working papers are disseminated for the purpose of generating comments and debate, and have not been subjected to peer review. The opinions expressed in this paper are exclusively those of the authors and do not necessarily represent the opinions of the RedNIE or its Board of Directors.

Citar como:
Brichetti, Juan Pablo, Eduardo Cavallo y Francisco Roch (2026). When Is the Sovereign Greenium a Scarcity Price? Green Supply, Market Depth, and Cross-Country Evidence. Documento de trabajo RedNIE N°413.

# When Is the Sovereign Greenium a Scarcity Price? Green Supply, Market Depth, and Cross-Country Evidence

Juan Pablo Brichetti  Eduardo A. Cavallo  Francisco Roch[1]

October 2026

We study how the sovereign greenium—the yield discount on green relative to conventional sovereign bonds—varies with green-bond supply and sovereign-market depth across countries. Using daily bond observations with week fixed effects on 104 sovereign green bonds in 29 countries over 2019–2025, we construct a bond-level synthetic greenium and document facts consistent with a relative-scarcity channel. A parsimonious multi-country framework shows how segmented demand for green assets, arising from investment mandates or investor preferences, can interact with limited and endogenous certified supply to generate a scarcity premium. Because certified supply responds to the greenium, the framework's negative supply prediction applies to observational regressions when variation is supply-driven. Lower sovereign green supply relative to the debt market is associated with higher greeniums: one percentage point of relative green supply is associated with a 2.5 basis point lower greenium. The negative association persists when relative supply is computed using a time-invariant, country-specific measure of market size. The relation is one between slow-moving market structures rather than a response to individual issues. Under the baseline time controls it is concentrated in emerging markets; the advanced-economy difference is sensitive to the treatment of common time shocks. Across countries, greeniums are higher in riskier sovereigns, and within countries the greenium compresses in acute sovereign stress, though both patterns rest almost entirely on Egypt. The highest-greenium sovereigns in the issue-size sample lack large green issues, a descriptive pattern consistent with thin green-market segments.

JEL: F34, G12, G15, H63, Q54.

Keywords: greenium, green bonds, sovereign debt, scarcity, investment mandates, sovereign risk.

## I. Introduction

Sovereign green bonds trade at a small yield discount relative to comparable conventional bonds. The average discount is a few basis points, but its dispersion across countries is wide: estimated greeniums range from negative values in some deep markets to close to a hundred basis points in stressed frontier issuers. Explanations centered on investor tastes for environmental assets account for the existence of a discount but have less to say about why it varies so much across sovereigns, why it moves over time within a sovereign, and why it is largest precisely where green issuance is thinnest.

This paper asks when the sovereign greenium behaves as the price of a scarce asset. A growing class of investors faces green investment mandates, regulatory or self-imposed requirements to hold labeled assets. Certified green supply, by contrast, is limited by the stock of projects that are financially

[1]Brichetti: Inter-American Development Bank, jbrichetti@iadb.org. Cavallo: Inter-American Development Bank, cavalloe@iadb.org. Roch: Universidad Torcuato Di Tella, francisco.roch@utdt.edu. The views expressed herein are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. All remaining errors are our own.

viable at the sovereign's cost of capital and worth certifying given their labeling costs. When mandated demand exceeds certified supply, investors bid up green bond prices, and the greenium emerges as the shadow price of the shortage. We formalize this logic in a deliberately simple multi-country framework and take its comparative statics to daily bond data with week fixed effects in 29 countries between 2019 and 2025.

The framework delivers a compact set of associations to look for in the data. Greeniums should be larger where certified green supply is small relative to the market in which mandated investors operate. They should be larger in countries with higher sovereign risk, because a higher cost of capital truncates the pool of financially viable projects and hence potential certified supply. And because certification costs per unit of capital are assumed to be higher for large, complex projects, smaller projects should enter the certified pool first, with larger projects appearing only where greeniums are unusually high. These are equilibrium cross-country predictions, and they carry two qualifications, each of which matters for interpreting one of our findings. The framework is static and has no state-contingent pricing. Certified supply, moreover, responds to the greenium itself, so the supply prediction holds in an observational regression only where variation is supply-driven — a condition Section II makes precise.

Measurement comes first, since greeniums are not observed. We construct a bond-level synthetic greenium by using each sovereign's conventional-bond curve to estimate the counterfactual spread of its green bonds. The resulting panel contains 72,609 bond-day observations on 104 green bonds in 29 countries. A central methodological finding is that the estimated level of the greenium is sensitive to the construction of the conventional counterfactual, while its slopes with respect to the main covariates are considerably more stable. A relabeling placebo likewise produces a small spurious greenium level but essentially none of the estimated slopes. We therefore organize the empirical analysis around covariate relationships rather than the precise average level of the greenium.

Three sets of findings emerge. First, sovereign greeniums covary negatively with relative green supply. A one-percentage-point increase in the outstanding stock of sovereign green bonds relative to the sovereign debt market is associated with a 2.45 basis point lower synthetic greenium. The sign and magnitude are stable across alternative constructions of the synthetic greenium, and the relabeling placebo does not reproduce the relationship. The association also persists when relative supply is computed using a time-invariant, country-specific denominator, indicating that it is carried by observed green-bond supply rather than by fluctuations in total sovereign debt. Inference is nevertheless sensitive to the limited number of country clusters, so we interpret the result as evidence of a stable reduced-form association rather than a precisely estimated causal effect.

The frequency of the identifying variation disciplines that interpretation. Country-year fixed effects substantially attenuate the coefficient, and country-quarter fixed effects drive it to zero. The association therefore reflects persistent differences in market structure rather than short-run responses to individual issues. At this frequency, the data cannot distinguish scarcity relief from slow-moving changes in green demand, market maturation, or broader financial development. Under the baseline time controls, the negative association is concentrated in emerging markets, while the advanced-economy slope is positive. The difference is sensitive to the treatment of common time shocks and may also be affected by imperfect measurement of tap issuance among frequent advanced-economy issuers. Different compo-

sitions of supply and demand shocks across market types provide one possible interpretation, but the paper does not identify those regimes.

Second, the sovereign-risk evidence is concentrated in a small number of stress episodes. Countries with higher average CDS spreads have higher average greeniums, while within countries the greenium compresses when sovereign risk rises. Both patterns, however, are driven largely by Egypt and lose significance when Egypt is excluded. The evidence therefore does not establish a general relationship between sovereign risk and the greenium. It instead documents that, in some episodes of acute stress, the pricing value of the green attribute compresses as default risk becomes dominant.

Third, sovereign issue size does not provide a valid test of the model's project-composition prediction. Sovereign green bonds finance portfolios of expenditures, so the size of an issue does not reveal the size or complexity of the underlying projects. Continuous measures of issue size are unrelated to country greeniums, while the absence of large green issues among the highest-greenium sovereigns is concentrated in a few thin markets. We interpret this as descriptive evidence on green-market depth, leaving the project-level composition prediction to be tested with use-of-proceeds data.

The paper makes three contributions. First, it develops a bond-level synthetic greenium and shows that its estimated level is sensitive to counterfactual construction, while its principal covariate relationships are more stable. Second, it documents a negative association between sovereign green-bond supply and the greenium in the pooled and emerging-market baseline specifications, and establishes that this relationship is not generated by fluctuations in the market-size denominator. Third, it identifies the frequency and scope conditions of that association: it is a feature of slow-moving market structures, is not robustly present in advanced economies, and cannot be separated from broader market development. The framework organizes these facts through the interaction of segmented green demand and costly, endogenous certified supply; the evidence supports that interpretation without identifying the origin of green demand or structurally validating the model.

*Related Literature*

The paper first speaks to the literature that measures the greenium. Corporate-market studies mostly find small average discounts, with magnitudes that vary widely across samples and methods (Zerbib, 2019; Larcker and Watts, 2020; Baker et al., 2022; Kapraun et al., 2021; Flammer, 2021; MacAskill et al., 2021; Gianfrate and Peri, 2019); sovereign estimates are fewer (Ando et al., 2022, 2024; Doronzo et al., 2021; Grzegorczyk and Wolff, 2022; D'Amico et al., 2023) and sensitive to sample, and our companion measurement work (Brichetti et al., 2025) documents wide cross-country dispersion in sovereign greeniums. Relative to this work we make two moves. We shift the object of interest from the average level to the cross-sectional and time-series structure of the greenium, and we show why the shift is necessary: the level of an estimated greenium is sensitive to the construction of the conventional counterfactual, to the point that reasonable specifications span the entire range from seven basis points to zero in our data, while slopes are not. That finding is itself a caution to the measurement literature, since part of the dispersion in published greenium estimates plausibly reflects counterfactual construction rather than pricing.

The mechanism places the paper in the literature on quantity-driven asset pricing in segmented markets. In preferred-habitat and demand-system frameworks, prices respond to supply and to the

portfolio demands of investor clienteles, not to a representative agent's preferences alone (Vayanos and Vila, 2021; Koijen and Yogo, 2019); bond supply moves risk premia when arbitrage capital is limited (Greenwood and Vayanos, 2014). Ours is a green-market instance of that logic. Mandated investors are the clientele, certified green bonds are the habitat, and the greenium prices the quantity mismatch. The nearest benchmark is the equilibrium taste model of Pástor et al. (2021), in which the green discount prices non-pecuniary preferences. The two views are not mutually exclusive: mandates and non-pecuniary preferences can both generate demand for green assets, and with market segmentation either can produce a scarcity premium when certified supply is limited. Our framework emphasizes this transmission from demand and costly certification to prices; the empirical analysis does not identify the origin of green demand.

Finally, the supply side of the framework draws on evidence about the costs of green labeling and verification, which microfound the certification threshold: issuance and certification carry fixed costs and reputational exposure (Ehlers and Packer, 2017; Caramichael and Rapp, 2022; Dyck et al., 2019; Flammer, 2013), with survey evidence more equivocal on how those costs scale with size (Sangiorgi and Schopohl, 2023; Bolibok, 2024). Our inference throughout follows the few-cluster literature (Cameron et al., 2008; Ibragimov and Müller, 2010), and the between-within decomposition follows Mundlak (1978).

Section II builds the framework and derives the condition under which its supply prediction maps into a regression. Section III constructs the synthetic greenium and shows by placebo which of its features can be trusted. Section IV contains the supply results: the baseline association, the alternative constructions of the supply ratio, and the frequency at which any of it is identified. Section V turns to sovereign risk, where one country carries the result, and Section VI to issue size. Section VII concludes.

## II. A Multi-Country Framework of Greenium Determination

### A. *Environment and Equilibrium Greenium*

There are $C$ countries. In country $c$, a representative investor with wealth $W_c$ allocates holdings between green bonds, which yield $i_{G,c} = r_f^c - g_c$, and conventional bonds, which yield $i_{SN,c} = r_f^c$, where $r_f^c$ is the local risk-free rate, interpreted as embedding sovereign and macroeconomic risk, and $g_c \geq 0$ is the greenium. A green mandate requires a fraction $m_c$ of the portfolio in green assets, generating mandated demand $M_c = m_c W_c$. The investor pays a quadratic penalty for shortfalls,

$$\max_{x_{G,c}} \left(r_f^c - g_c\right) x_{G,c} + r_f^c \left(W_c - x_{G,c}\right) - \frac{\phi_c}{2} \left(M_c - x_{G,c}\right)^2, \tag{1}$$

where $\phi_c > 0$ indexes the strictness of the mandate. The first-order condition and market clearing ($x_{G,c} = S_{G,c}$, the certified green supply) give

$$g_c = \phi_c \left(M_c - S_{G,c}\right). \tag{2}$$

The greenium is the shadow price of the shortage of certified green assets relative to mandated demand: zero when supply meets the mandate, positive and increasing in the gap otherwise, with $\phi_c$ amplifying

the sensitivity. Throughout, $g_c$ is the scarcity component of the observed greenium; other pricing forces and measurement error are left outside the framework, and Section III returns to the distinction. Pricing is by a predominantly market-specific marginal investor, a reduced form for jurisdiction-specific regulation, benchmark and hedging frictions, and certification credibility differences that make cross-border substitution imperfect; this is what allows local scarcity to sustain cross-country dispersion in greeniums.

Because $M_c = m_c W_c$ scales with the financial system, cross-country comparisons are naturally conducted in ratios. Under the normalization $\phi_c = \tilde{\phi}/W_c$, which makes the penalty comparable across market sizes, equation (2) becomes

$$g_c = \tilde{\phi}\left(m_c - \frac{S_{G,c}}{W_c}\right), \tag{3}$$

so the greenium is decreasing in *relative* green supply $s_c = S_{G,c}/W_c$ holding the mandate intensity $m_c$ fixed. The normalization is not innocuous, since it converts the level gap $M_c - S_{G,c}$ into a testable ratio. It is exactly what obtains if the mandate penalty is quadratic in the portfolio-*share* shortfall, $\frac{\tilde{\phi}}{2} W_c \,(m_c - x_{G,c}/W_c)^2$, which is how prudential and benchmark-tracking mandates are written in practice, as fractions of assets rather than currency amounts. Without it, the model's empirical object would be the level gap, which is not comparable across a sample whose market sizes span three orders of magnitude. The ratio $s_c$ is the empirical object in Section IV.

### *B. Viable Projects, Certification, and Green Supply*

Green supply is built from projects. Country $c$ has potential masses $\bar{S}_{j,c}$ of projects of two sizes $k_1 < k_2$, with gross returns $R_{j,c}$ drawn from country-specific distributions. A project is financially viable at the conventional rate if $R_{j,c} \geq r_f^c$, so the viable mass is $S_{j,c}^{SN} = \nu_{j,c}(r_f^c)\,\bar{S}_{j,c}$ with $\nu_{j,c}(r) = 1 - F_{j,c}(r)$ decreasing in $r$: a higher cost of capital truncates lower-return projects.

Certification carries a fixed cost $\kappa_{j,c}$, with $\kappa_{2,c} > \kappa_{1,c}$ and $\kappa_{2,c}/k_2 > \kappa_{1,c}/k_1$: large, complex projects are costlier to certify both absolutely and per unit of capital, reflecting broader environmental assessment, more intensive monitoring, and reputational exposure that plausibly rises with scale and visibility (Caramichael and Rapp, 2022; Bolibok, 2024; Dyck et al., 2019; Flammer, 2013); survey evidence is more equivocal on the size gradient (Sangiorgi and Schopohl, 2023). A viable project certifies when the interest saving covers the cost, $g_c k_j \geq \kappa_{j,c}$, defining thresholds $g_{j,c}^\star = \kappa_{j,c}/k_j$ with $g_{2,c}^\star > g_{1,c}^\star$. Certified supply is then a step function,

$$S_{G,c}(g_c) = S_{1,c}^{SN}\mathbf{1}\{g_c \geq g_{1,c}^\star\} + S_{2,c}^{SN}\mathbf{1}\{g_c \geq g_{2,c}^\star\}, \tag{4}$$

We read equation (4) as a correspondence at the certification thresholds. At $g_c = g_{j,c}^\star$ the marginal project is indifferent, so any fraction of the indifferent mass may certify, and certified supply takes any value between the mass already certified and that mass augmented by $S_{j,c}^{SN}$. The equilibrium greenium then solves $g_c^* = \phi_c\big(M_c - S_{G,c}(g_c^*)\big)$. Since the right-hand side is weakly decreasing in $g_c$ and the left-hand side increases, the equilibrium exists and is unique. Without the indifference margin the all-or-nothing rule can leave no fixed point, because neither certifying nothing nor certifying an entire mass need be consistent with the greenium each would imply.

### C. Predictions

The comparative statics we take to the data are the following. Holding mandates fixed, the greenium is decreasing in relative green supply $s_c = S_{G,c}/W_c$, from equation (3). Holding potential supply fixed, a higher $r_f^c$ reduces $\nu_{j,c}$, shrinks certified supply, and raises the greenium, so riskier sovereigns should display higher greeniums in the cross-section. And since small projects clear the certification hurdle first, green segments should be dominated by small projects unless the greenium exceeds the upper threshold $g^{\star}_{2,c}$, so large certified projects should appear only where greeniums are unusually high.

A remark disciplines the mapping from the first prediction to a regression of $g$ on $s$, and matters for everything in Section IV. The prediction $\partial g/\partial s < 0$ holds mandated demand fixed, but certified supply in this framework is not exogenous: it responds to the greenium through the certification thresholds of equation (4). Consider the equilibrium $g = \tilde{\phi}\,[m - s(g,z)]$, where $s(\cdot)$ is the certification schedule, weakly increasing in $g$, and $z$ shifts the viable project pool. A supply shock ($dz > 0$ at fixed $m$) moves $g$ and $s$ in opposite directions; a mandate shock ($dm > 0$) raises $g$ and, where thresholds are crossed, raises $s$ as well, so the two co-move positively. Because supply moves in steps, this is a statement about signs of co-movement, not about a smooth slope. Under simplifying assumptions, a common $\tilde{\phi}$, no other shocks correlated with $s$, and abstracting from the fixed effects that residualize the estimating equation, the population slope of $g$ on $s$ illustrates the same point: $\beta = \tilde{\phi}\,[\mathrm{Cov}(m,s)/\,\mathrm{Var}(s) - 1]$, which equals $-\tilde{\phi}$ only when supply variation is orthogonal to mandates and turns positive when demand-induced issuance dominates. The formula illustrates the point; it is not the exact mapping to our regressions. The condition also gives the regression coefficient a structural unit. Where variation in $s$ is supply-driven, the population slope is $-\tilde{\phi}$: the magnitude of the coefficient measures the sensitivity of the greenium to relative supply, in basis points per percentage point. Because the empirical analysis does not identify whether the observed variation in $s$ is supply-driven and orthogonal to mandates, however, we do not interpret the estimated coefficient as a structural estimate of $\widetilde{\phi}$. The same pricing condition follows from a non-pecuniary holding benefit $a_c x_{G,c} - (b_c/2)x_{G,c}^2$ with $a_c = \phi_c M_c$ and $b_c = \phi_c$. Mandates and preferences are thus alternative foundations for green demand; the scarcity mechanism concerns how that demand meets limited certified supply.

The empirical prediction we take to Section IV is accordingly conditional, and stating it this way preserves what the data can reject. The negative partial derivative is a prediction of the model conditional on supply-driven variation, so a positive observational slope rejects the unconditional empirical prediction as we have stated it; it need not reject the broader equilibrium mechanism, which can generate a positive slope when supply responds endogenously to demand shocks. Distinguishing the two requires evidence on the composition of shocks, which we do not have, so a sign reversal is a finding to be explained rather than a result the framework accommodates by construction.

Two comments delimit what the framework does and does not predict. The mandate-intensity prediction, that greeniums rise with $m_c$ or with mandated demand relative to market size, requires data on mandates, which we do not observe; it remains untested here. And the framework is a static equilibrium: it compares steady states across countries and says nothing about how the greenium responds to a sovereign risk shock within a country over weeks, or at what frequency the supply comparative static should be visible. The distinction between persistent cross-country structure and high-frequency dynamics turns out to be exactly where the data speak.

## III. Data and Measurement

### A. Bond Panel

The bond data are daily z-spreads, tenors, and bid-ask spreads for sovereign green and conventional bonds from Refinitiv/LSEG over 2016–2025, the same raw source as our companion measurement work (Brichetti et al., 2025); the estimation samples used here effectively span 2019–2025, the period with meaningful sovereign green issuance. Although bond observations are daily, the country-level supply variables, lags, and time fixed effects are defined by week; the baseline therefore retains daily bond observations with week fixed effects, and Section IV reports a bond-week collapse as a robustness check. The green universe underlying the supply measure comprises 206 sovereign green bonds (175 with known face value); the synthetic greenium panel described below retains 104 green bonds in 29 countries, the subset with comparable conventional bonds and the covariates required by the first stage. Sample sizes vary slightly across exercises because the risk and matching panels impose different data requirements; each table reports its own coverage, and the appendix reconciles panel composition country by country.

Two country-level series complete the data. Relative green supply is $s_{c,t} = S_{G,c}(t)/W_c(t)$. The numerator $S_{G,c}(t)$ is the outstanding stock of sovereign green bonds by country and week, built from issuance and maturity dates, using amounts outstanding rather than issued for frequent tappers such as Austria, France, Germany, Belgium, and Ireland. The denominator $W_c(t)$ is the BIS stock of general government debt securities, quarterly, carried forward to weeks. We use it as an observable proxy for the scale variable in Section II, not as a measure of portfolio wealth subject to green mandates. The ratio describes the green share of sovereign debt; mandated demand remains unobserved. Because the BIS carry-forward changes at week rather than month boundaries, quarters in Section IV.E are defined by the assigned BIS quarter, within which $W_c$ is constant in all 487 country-quarter cells. The ratio is available for 30 of 40 countries; Egypt is among those missing, for lack of a BIS market-size series, a fact that matters when interpreting Section V. Country averages $\bar{s}_c$ range from 0.01 percent (China) to 15.1 percent (Chile) with a median of 2.6 percent, a standard deviation of 3.75 percentage points across country-weeks, and a P10–P90 span of 0.35 to 9.25 percent. We measure sovereign risk with the five-year sovereign CDS spread and, alternatively, with the I-spread from the IMF's Sovereign Spread Monitor, an internal IMF database. Known measurement limitations are worth stating: 15 percent of green bonds lack face value; tap timing is imperfectly captured; $W_c$ does not separate domestic from foreign holdings and excludes corporate and supranational green substitutes; and its quarterly frequency means within-quarter variation in $s_{c,t}$ comes from the numerator.

### B. The Synthetic Greenium

The greenium is estimated bond by bond in two stages. For each country we fit, on conventional bonds only,

$$\text{zspread}_{j,t}^{brown} = \gamma_1 \,\text{tenor}_{j,t} + \gamma_2 \,\text{bid-ask}_{j,t} + \delta_{w(t)} \,[+\varphi_k] + u_{j,t}, \tag{5}$$

with week fixed effects $\delta_{w(t)}$ and currency fixed effects $\varphi_k$ where the country issues in several currencies, with $w(t)$ denoting the week containing date $t$. We then predict the counterfactual z-spread of

each green bond $i$ on date $t$ and define the synthetic greenium as $\hat{g}_{i,t} = \widehat{\text{zspread}}_{i,t}^{cf} - \text{zspread}_{i,t}^{obs}$, positive when the green bond trades tighter than a comparable conventional. The construction avoids tying the greenium to whichever conventional bonds happen to quote on a given date, and it supports bond fixed effects in second-stage regressions. First-stage fit is good: the median $R^2$ across 37 country regressions is 0.836, only China falls below 0.50 (0.12, 0.2 percent of the panel), and flexible tenor terms raise the median above 0.90 without changing any second-stage slope materially (appendix Table B10 reports the full distribution). The synthetic greenium is itself an estimate, and the second-stage standard errors we report treat it as data, ignoring the sampling error in equation (5); they therefore understate uncertainty by ignoring the sampling error of the first-stage curve. Classical noise in $\hat{g}$ does not bias the slopes, but first-stage error correlated with the second-stage regressors could bias them in either direction; the relabeling placebo of the next subsection is the empirical check that it does not.

The resulting panel has 72,609 bond-day observations, corresponding to 15,087 unique bond-week cells, 104 green bonds, and 29 countries (Belgium drops for lack of comparable conventionals); appendix Table B1 reconciles it with the risk panels. The mean synthetic greenium is $+5.42$ basis points and the median $+3.18$, confirming that sovereign green bonds trade at a modest average discount, with wide cross-country heterogeneity, from negative greeniums in Saudi Arabia and Korea to clearly positive ones in Iceland and the Philippines. The framework restricts the scarcity component of the greenium to be non-negative; the estimated greenium also contains measurement error and whatever else prices green bonds, so negative estimates are not by themselves evidence against it. Two robustness constructions are used throughout: a direct cell-matching greenium against conventional bonds in the same country-currency-maturity cell, and a country-matching greenium against the country's average conventional bond.

### *C. Levels versus Slopes*

Part of the measured greenium could be manufactured by the construction itself. A placebo bounds how much: in each country we randomly relabel up to a third of conventional bonds as green, exclude them from first-stage estimation, and compute their placebo greenium, over 200 iterations. The placebo mean is $+1.60$ basis points (95 percent confidence interval $[1.28, 1.91]$ across iterations), about 22 percent of the real mean of $+7.19$ in the corresponding panel. The level of the synthetic greenium therefore carries a mechanical positive bias. Consistent with this, the mean level is highly sensitive to the first stage, moving from $+7.19$ (baseline) to $+1.60$ under same-currency estimation with quadratic tenor and to $-0.06$ basis points under same-currency estimation with a maturity spline; appendix Tables B2 and B9 report the variants and the placebo distributions in full.

Slopes behave differently. Re-estimating the second-stage regressions of Sections IV and V on each placebo iteration yields slope distributions centered on zero: the placebo scarcity slope has mean $-0.13$ and standard deviation 0.97 against a real slope of $-2.45$, which sits at the first percentile (pseudo $p = 0.015$); the placebo CDS slope has mean $+0.43$ and standard deviation 9.94 against a real slope of $-19.70$, at the second percentile (pseudo $p = 0.045$), the fatter tail reflecting relabeled bonds in high-CDS-variance countries. The first-stage bias is a level bias, not a covariance bias. Slope estimates are correspondingly stable across first-stage variants, within $[-2.45, -1.64]$ for scarcity and $[-22.9, -19.7]$ for CDS; Figure 1 displays the contrast, with the mean greenium collapsing from $+7.2$

basis points to zero across the same five constructions under which the risk slope barely moves. All conclusions in this paper rest on slopes; we make no claim about the level of the greenium beyond its order of magnitude.

## IV. Green Supply Scarcity and the Greenium

### A. *Baseline Association and Inference*

The framework's central comparative static links the greenium to relative green supply. The second-stage regression is

$$\hat{g}_{i,t} = \delta\, s_{c,t} + \alpha_i + \gamma_{w(t)} + \varepsilon_{i,t}, \tag{6}$$

with bond fixed effects $\alpha_i$ and week fixed effects $\gamma_{w(t)}$. The model's supply comparative static predicts $\delta < 0$ when the identifying variation is supply-driven. Table 1 reports the estimates.

The baseline coefficient is $\hat{\delta} = -2.45$: a one percentage point increase in relative green supply is associated with a 2.45 basis point lower synthetic greenium. Collapsing the daily panel to one observation per bond-week leaves the estimate virtually unchanged at $-2.420$ ($N = 15{,}087$). Scaled by the dispersion of $s$ across country-weeks, a one standard deviation movement (3.75 percentage points) corresponds to roughly 9.2 basis points, and the P10–P90 span to roughly 21.8 basis points. These magnitudes compare distribution extremes across countries and weeks, not a typical within-country shock, and the synthetic greenium's high cross-bond variance is why the implied span can exceed the mean greenium; the dummy-variable specification of appendix Table B6 brackets the association from beneath with a much smaller magnitude. The coefficient is similar with $s$ lagged four and thirteen weeks ($-2.39$ and $-2.10$). This is not evidence on causality: Section IV.E shows the identifying variation is slower than any such lag window. The association is concentrated in emerging markets ($-2.93$); in advanced economies the sign reverses ($+0.97$), a result we take up below. Excluding the

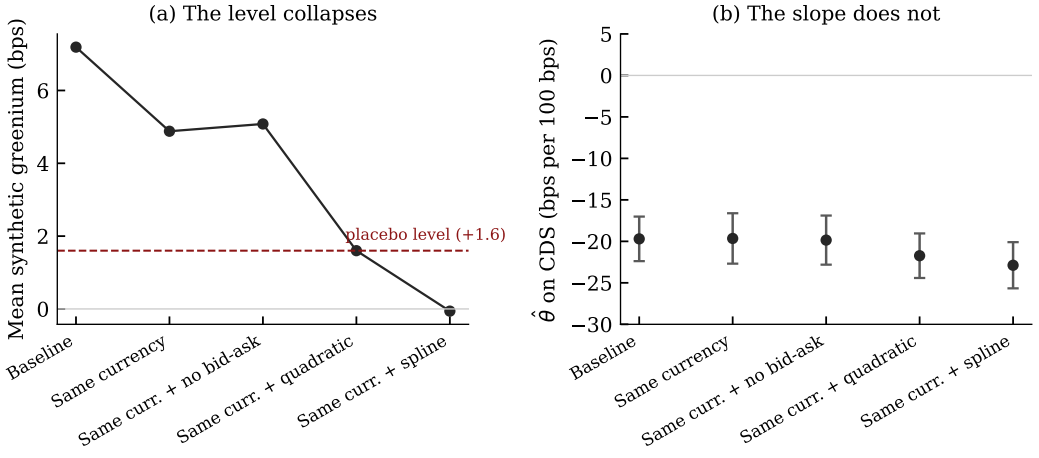


FIGURE 1. The level is fragile; the slope is not. Panel (a): mean synthetic greenium across five first-stage constructions, against the placebo level of +1.6 basis points. Panel (b): the within-country slope of the greenium on CDS across the same constructions, with 95 percent confidence intervals clustered by country and week.

TABLE 1. Relative Green Supply and the Synthetic Greenium

| | (1) Levels | (2) Log | (3) EMDE | (4) AE | (5) Lag 4w | (6) Lag 13w |
|---|---|---|---|---|---|---|
| Relative green supply $s_{c,t}$ | -2.45** | -3.41 | -2.93** | 0.97*** | -2.39** | -2.10** |
| | (0.15) | (0.49) | (0.19) | (0.09) | (0.13) | (0.15) |
| | [1.03] | [2.71] | [1.17] | [0.30] | [1.02] | [0.93] |
| Observations | 72,609 | 72,609 | 49,817 | 22,792 | 72,050 | 70,586 |
| $R^2$ | 0.407 | 0.397 | 0.406 | 0.617 | 0.407 | 0.408 |
| $p$, two-way cluster | 0.024 | 0.219 | 0.023 | 0.009 | 0.027 | 0.032 |
| $p$, wild cluster bootstrap | 0.156 | 0.319 | 0.190 | 0.029 | 0.164 | 0.173 |

*Notes.* Dependent variable: synthetic greenium (bps) at the bond–day level. All columns include bond and week fixed effects. Column (2) replaces $s_{c,t}$ with $\log s_{c,t}$; columns (5)–(6) use $s$ lagged four and thirteen weeks. Standard errors clustered by year–week in parentheses; two-way (country×week) in brackets. Wild cluster restricted bootstrap: 999 Rademacher replications, clustered by country (29/18/11 clusters depending on the sample). Stars follow two-way clustering, the country-robust inference the paper privileges; week-clustered stars would overstate significance with 29 country clusters. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

three countries with the lowest average $s$ (China, Korea, Indonesia) leaves the coefficient unchanged at $-2.45$, and winsorizing $s$ at the 1st and 99th percentiles moves it to $-2.63$.

Figure 2 summarizes how the slope behaves across constructions and inference strategies. The upper block re-estimates equation (6) under the baseline first stage and four variants: dropping the bid-ask control, quadratic tenor, a natural tenor spline, and estimation within country-currency blocks. The coefficient stays negative and significant under two-way clustering throughout, between $-1.64$ and $-2.45$, with the same-currency variant the most precise ($p_{tw} = 0.003$); the scarcity slope is not an artifact of the conventional-curve specification.

The lower block varies the inference. With roughly 452 week clusters the estimate is sharp; the binding constraint is the country dimension, with 29 clusters. Two-way country-week clustering gives $p = 0.024$. A wild cluster restricted bootstrap on countries (Cameron et al., 2008) gives $p = 0.156$. Ibragimov–Müller (Ibragimov and Müller, 2010), which estimates the slope country by country and tests the mean across the 16 countries with valid estimates, gives $p = 0.275$, and the leave-one-country-out jackknife $p = 0.360$. Table 2 collects the methods. The pattern is what one expects when a moderate signal meets few effective clusters: strategies that weight countries equally lose power, while strategies that preserve the pooled weighting retain it. The country-level inference is suggestive. The slope placebo of Section III.C is the sharpest evidence that the association is not mechanical: constructions with no green content essentially never reproduce it.

### *B. Green Supply and Market Size*

To examine whether the baseline association depends on the construction of the supply ratio, we compare it with measures that hold either component at its country mean. The exercise uses the synthetic panel, the baseline fixed effects, and the same inference methods.

Entering $\log S_{G,c}$ and $\log W_c$ jointly is uninformative: the two series are collinear enough within countries that no component is significant in any group, and standard errors exceed the coefficients

TABLE 2. Alternative Inference with Few Country Clusters

| *Panel A: scarcity slope (synthetic panel, 29 country clusters)* Method | $\hat{\delta}$ | SE | $t$ | $p$ |
|---|---|---|---|---|
| Pooled | -2.45 | 0.15 | -15.96 | 0.000 |
| Cluster: week | -2.45 | 0.15 | -15.96 | 0.000 |
| Two-way (country x week) | -2.45 | 1.03 | -2.38 | 0.024 |
| Wild cluster bootstrap | -2.45 | | -2.37 | 0.156 |
| Ibragimov-Muller | -2.43 | 2.15 | -1.13 | 0.275 |
| Jackknife (leave-one-country-out) | -2.45 | 2.63 | -0.93 | 0.360 |
| *Panel B: sovereign-risk slope* Panel / risk / method | $\hat{\theta}$ | SE | $t$ | $p$ |
| Synthetic, CDS: Wild cluster bootstrap | -19.70 | | | 0.0003 |
| Synthetic, I-Spread: Wild cluster bootstrap | -30.40 | | | 0.0016 |
| Cell, CDS: Wild cluster bootstrap | -4.97 | | | 0.3574 |
| Cell, I-Spread: Wild cluster bootstrap | -7.02 | | | 0.6518 |
| Country, CDS: Wild cluster bootstrap | -19.40 | | | 0.0155 |
| Country, I-Spread: Wild cluster bootstrap | -28.52 | | | 0.1031 |
| Synthetic, CDS: Ibragimov-Muller | -6.67 | 4.27 | -1.56 | 0.1277 |
| Synthetic, CDS: Jackknife LOO | -19.70 | 14.08 | -1.40 | 0.1707 |
| Synthetic, I-Spread: Ibragimov-Muller | -3.05 | 2.79 | -1.10 | 0.2839 |
| Synthetic, I-Spread: Jackknife LOO | -30.40 | 21.85 | -1.39 | 0.1765 |

*Notes.* Wild cluster restricted bootstrap: Rademacher weights, clustered by country, 999 (Panel A) and 9,999 (Panel B) replications, imposing the null. Ibragimov–Müller estimates country-level slopes (bond fixed effects, no week effects within country) and applies a $t$-test across countries; only 16 (Panel A) and 33/26 (Panel B) countries yield valid estimates. The jackknife re-estimates the pooled regression leaving one country out at a time.

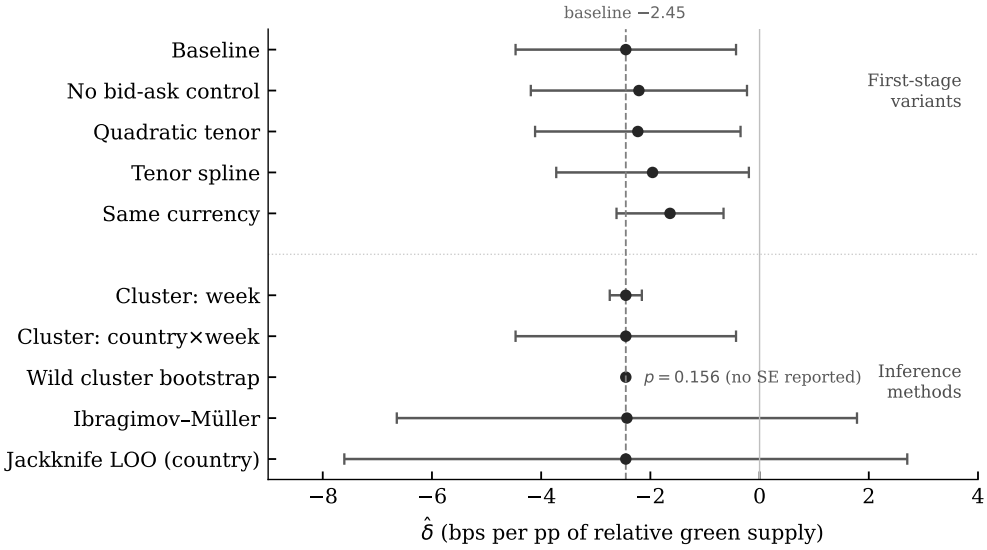


FIGURE 2. Robustness of the scarcity slope. The upper block re-estimates the baseline regression under five first-stage constructions of the synthetic greenium, with 95 percent two-way-clustered confidence intervals. The lower block holds the baseline estimate fixed and varies the inference method; the wild cluster bootstrap reports a $p$-value only. The dashed line marks the baseline $\hat{\delta} = -2.45$.

(appendix Table B4). We therefore consider two alternative constructions on the scale of $s$: the fixed-denominator ratio $s^{num}_{c,t} = S_{G,c}(t)/\bar{W}_c$ and the fixed-numerator ratio $s^{den}_{c,t} = \bar{S}_{G,c}/W_c(t)$. Table 3 reports the results.

With the fixed-denominator ratio, the negative association is slightly larger in magnitude than in the baseline: $-2.88$ pooled ($p = 0.015$) and $-3.34$ in emerging markets ($p = 0.011$). The pooled coefficient remains negative in every leave-one-country-out estimate. This supports the stability of the

TABLE 3. Green Supply versus Market Size: Counterfactual Ratios

| | Ratio $s$ | Numerator only $s^{num}$ | Denominator only $s^{den}$ |
|---|---|---|---|
| Pooled (N=72,490) | -2.445** (1.03; $p$ 0.025) | -2.877** (1.12; $p$ 0.015) | +2.374** (1.05; $p$ 0.031) |
| EMDE (N=49,747) | -2.926** (1.17; $p$ 0.023) | -3.338** (1.18; $p$ 0.011) | +3.620* (1.71; $p$ 0.05) |
| AE (N=22,743) | +0.975*** (0.31; $p$ 0.01) | +0.933*** (0.29; $p$ 0.01) | +1.774 (1.17; $p$ 0.159) |

*Notes.* Dependent variable: synthetic greenium (bps) at the bond–day level, with bond and week fixed effects; two-way (country×week) standard errors in parentheses. $s^{num}_{c,t} = S_{G,c}(t)/\bar{W}_c \times 100$ lets only green supply move, holding the denominator at its country mean; $s^{den}_{c,t} = \bar{S}_{G,c}/W_c(t) \times 100$ lets only market size move. Both are on the scale of $s$ ($s^{num}$ matches its mean, 5.65 vs. 5.66; $s^{den}$ has mean 5.74). The joint specification in logs is uninformative by collinearity (appendix Table B4). The EMDE denominator coefficient sits at the 5 percent boundary ($p = 0.0498$). The two counterfactual ratios are distinct regressors, not an algebraic decomposition: their coefficients do not sum to the baseline estimate. Leave-one-country-out: the ratio slope spans $[-4.93, -1.63]$ and stays significant at 5 percent in 28 of 29 exclusions, the numerator $[-6.40, -2.06]$ (28/29), the denominator $[+1.72, +5.83]$ (26/29); Chile is the most influential country and its exclusion strengthens both coefficients. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

association to removing time variation in the denominator. The fixed-numerator coefficient is $+2.37$ pooled ($p = 0.031$); the emerging-market estimate lies at the 5 percent significance boundary. Appendix Table B5 reports a related exercise on the auxiliary green-dummy panel, where neither interaction is significant under two-way clustering when green supply and market size enter jointly.

These ratios are alternative regressors, not an algebraic or economic decomposition. Holding a component fixed in the regressor does not hold it fixed in the underlying market: green supply and market size continue to move jointly. A positive coefficient on $s^{den}$ can therefore arise even under a negative relationship between the greenium and the actual ratio $s$; it neither refutes a scarcity mechanism nor identifies a separate market-depth effect. The central result is that the negative pooled and emerging-market associations do not require fluctuations in the denominator. Section IV.E examines the frequency of the variation supporting that result.

### *C. Nonlinearity*

The average gradient may mask heterogeneity across the distribution of relative supply. We examine that with a three-piece linear spline in $s$, with knots at the 20th and 80th percentiles of the bond–day distribution, bond and week fixed effects, and two-way standard errors. Appendix Table B3 reports it.

The middle of the distribution carries the association, and carries it more steeply than the pooled estimate: $-4.68$ between the 20th and 80th percentiles ($p = 0.027$, 22 countries). Above the 80th percentile the slope attenuates to $-1.08$ ($p = 0.005$). Below the 20th percentile the sign goes the other way, $+10.59$, though imprecisely ($p = 0.253$); that segment collects observations with almost no green supply, among them China, whose first-stage fit is the weakest in the panel.

Several limitations bound this reading. The framework does not predict where or whether the slope should change, so the exercise is descriptive and the knots carry no structural interpretation. We have not tested whether the segment slopes differ from one another, and that one is significant and another is not says nothing about the difference between them. The top segment rests on four countries. The pattern does not survive the alternative green-dummy construction, where the segment interactions are ordered differently and none is significant under two-way clustering. And a slowly maturing market can produce a profile of this shape on its own, a possibility Section IV.E makes concrete. We read the spline as evidence that the average gradient masks heterogeneity across the distribution of relative supply, and not as a test of the framework.

### *D. Advanced and Emerging Markets*

Under the country-level inference this paper privileges, the only slope in Table 1 that rejects is the advanced-economy one, and its sign is the opposite of the prediction ($+0.97$, $p_{tw} = 0.009$, bootstrap $p = 0.029$), while the pooled and emerging-market estimates do not clear the bootstrap. The difference between the two slopes is $+3.90$, with two-way $p = 0.003$ and a wild cluster bootstrap $p$ of about 0.03 under the baseline group-specific time controls. A level difference in the same direction is already documented for sovereign greeniums (Ando et al., 2024); what the split here concerns is the gradient with respect to relative supply, not the average level. The difference is sensitive to the treatment of common time shocks: with week fixed effects common to both groups the advanced-economy slope is no longer positive ($-1.35$) and the difference falls to $+1.20$ and loses significance, and with country-

quarter fixed effects it vanishes, placing the heterogeneity at the same low frequency as the baseline association (Table 4). Held to a uniform standard, the sharpest statement the country-robust inference supports is therefore that supply and the greenium co-move *positively* in advanced economies, and negatively, with suggestive country-level significance, in emerging markets. The remark of Section II.C says what a sign reversal can and cannot mean: because certified supply responds to the greenium, the regression sign can reflect the composition of shocks, so the positive advanced-economy slope rejects the unconditional prediction as stated but need not reject the equilibrium mechanism if the observed supply movements there are demand-induced. The positive advanced-economy coefficient also appears with the fixed-denominator ratio ($+0.93$, $p = 0.010$; Section IV.B). Both specifications use the same green-supply series, so imperfect measurement of tap timing among frequent advanced-economy issuers could contribute to the split. One possible interpretation is that the composition of shocks differs across market types, with supply-driven variation dominating in one group and demand-induced issuance in the other. We do not observe which shocks dominate. Distinguishing the two regimes requires evidence on whether issuance responds differently to lagged greeniums across the two groups. The negative association observed in emerging markets is not robustly established in advanced economies.

### *E. The Frequency of Identifying Variation*

At what frequency is the association identified? The question disciplines every interpretation above, and we answer it by adding progressively finer country-time fixed effects to the baseline regression. Table 5 reports the ladder.

With country-year fixed effects the pooled coefficient falls from $-2.45$ to $-0.72$ ($p = 0.12$); with country-quarter fixed effects, which absorb the quarterly $W_c$ mechanically and leave only intra-quarter issuance and maturity variation in $S_{G,c}$, it is $-0.08$ ($p = 0.94$). The same happens in each group, to the numerator-only ratio, and to the difference between advanced and emerging economies. After residualizing $s$ on bond and week fixed effects, adding country-quarter fixed effects leaves only 5.3 percent of the baseline residual variance; the remaining 94.7 percent is absorbed. In addition, 405 of 487 BIS-assigned country-quarter cells (83.2 percent; 78.3 percent when weighted by daily observations) contain no intra-quarter variation. The standard error is nevertheless essentially unchanged (0.98

TABLE 4. Equality of Scarcity Slopes: Advanced versus Emerging Economies

| Specification | EMDE | AE | Difference | $p$, two-way | $p$, bootstrap |
|---|---|---|---|---|---|
| Group-specific week FE | -2.926 | +0.975 | +3.900 | 0.003 | 0.029 |
| Common week FE | -2.551 | -1.351 | +1.201 | 0.319 | |
| Common week FE + country-quarter FE | -0.013 | -0.066 | -0.053 | 0.966 | |

*Notes.* Formal test of equality of the scarcity slope across advanced and emerging economies. The first row replicates the group-specific week fixed effects of Table 1 (difference standard error 1.202, groups disjoint). Wild cluster restricted bootstrap: 999 Rademacher replications with the null imposed, 29 country clusters; stable across five seeds ($p$ between 0.024 and 0.038, mean 0.029). With week fixed effects common to both groups the difference falls to $+1.20$ and loses significance; with country–quarter fixed effects it vanishes, so the heterogeneity lives at the same low frequency as the baseline association; note that under common week fixed effects the advanced-economy slope is no longer positive. $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.10$.

TABLE 5. Where the Variation Lives: Country–Time Fixed Effects

| | Bond + week FE | + country–year FE | + country–quarter FE |
|---|---|---|---|
| Pooled | -2.445** (1.03) | -0.715 (0.45; $p$ 0.12) | -0.080 (0.98; $p$ 0.94) |
| EMDE | -2.926** (1.17) | -0.780 (0.54; $p$ 0.17) | -0.013 (1.21; $p$ 0.99) |
| AE | +0.975*** (0.31) | +0.185 (0.29; $p$ 0.54) | -0.066 (0.32; $p$ 0.84) |

*Notes.* Coefficient on relative green supply $s_{c,t}$, dependent variable the synthetic greenium, adding progressively finer country–time fixed effects; two-way standard errors. Quarters in the last column are assigned BIS quarters, within which $W_c$ is constant in all 487 country–quarter cells. Under calendar-quarter grouping, $W_c$ is constant in 112 of 489 cells (22.9 percent). Country–quarter fixed effects therefore absorb $W_c$ mechanically, so the last column speaks only to the numerator margin. After residualizing $s$ on bond and week fixed effects, adding country–quarter fixed effects leaves 5.3 percent of the baseline residual variance; this calculation uses the 72,609 daily observations with equal observation weights. Moreover, 405 of 487 BIS-assigned country–quarter cells (83.2 percent; 78.3 percent when weighted by daily observations) contain no intra-quarter variation. The standard error is essentially unchanged (0.98 vs. 1.03), and the 95 percent interval $[-2.09, +1.93]$ excludes the baseline $-2.45$, so the attenuation reflects primarily a change in the estimated relationship rather than a loss of precision; the difference between the two estimates is not formally tested. The numerator-only ratio behaves identically ($+0.10$ pooled, $+0.21$ EMDE with country–quarter FE), and with country–year FE the numerator falls to $-1.11$ ($p = 0.12$) and the denominator to $+0.68$ ($p = 0.44$). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

against 1.03), and the 95 percent interval, $[-2.09, +1.93]$, excludes the baseline estimate. We do not formally test the difference between the two estimates, which are computed on overlapping variation. Because $s^{den}$ is constant within assigned BIS quarters, the country-quarter test speaks only to the numerator margin; the country-year result, which still leaves quarterly variation in both margins, shows that neither survives a flexible country trend.

The finding sets the terms for interpretation, in both directions. It is not inconsistent with the framework, which is a static comparison across persistent market structures rather than a theory of short-run dynamics, but the framework places no restriction on frequency and cannot be said to anticipate the result. Two inferences that the high frequency of the bond panel invites do not survive, however. Lagging supply by four or thirteen weeks provides no protection against reverse causality when the identifying variation moves at horizons beyond a year. And the contrast between the within-country margin and the cross-section is weaker than a high-frequency panel suggests: bond fixed effects absorb time-invariant country characteristics, but slow-moving ones, such as mandate growth, market maturation, and financial development, are exactly what below-annual variation cannot exclude. These limits concern what the evidence identifies, not the economic logic that organizes it. Investment mandates give green demand one institutional interpretation, but non-pecuniary preferences can generate the same pricing condition. The framework links this demand to project viability and costly certification, showing how limited certified supply can generate a scarcity premium under either interpretation. The pooled and emerging-market estimates have the sign of the model's supply comparative static. The negative pooled association persists across the first-stage specifications considered and when relative supply is computed with a time-invariant, country-specific denominator; slopes as negative as the baseline are

rare in the relabeling placebo. These findings are consistent with a scarcity channel but do not establish that the supply-demand gap causes the association, which may also reflect slow-moving changes in green demand or broader market development.

## V. Sovereign Risk and the Greenium

### A. *Between and Within*

The framework predicts that riskier sovereigns have scarcer certifiable supply and hence higher greeniums. This is a cross-country, equilibrium prediction. The within-bond regression of the synthetic greenium on the five-year CDS spread,

$$\hat{g}_{i,t} = \theta \, \text{Risk}_{c(i),t} + \alpha_i + \gamma_{w(t)} + \varepsilon_{i,t}, \tag{7}$$

identifies instead the response to risk *movements* within a sovereign. Table 6 reports it, and Table 7 separates the two margins with a Mundlak decomposition (Mundlak, 1978).

The between coefficient is positive: countries with 100 basis points higher average CDS have synthetic greeniums about 8.8 basis points higher on average (+15.7 under country-matching), and the country-level scatter has a correlation of 0.59. The direction matches the framework's cross-sectional prediction, but the estimate does not survive Egypt: excluding it, the between coefficient flips to −3.16

Table 6. Sovereign Risk and the Synthetic Greenium (Within Variation)

| | Sovereign CDS, 5-year | | | IMF I-Spread | | |
|---|---|---|---|---|---|---|
| | All | Excl. Egypt | EMDE | All | Excl. Egypt | EMDE |
| Sovereign risk (per 100 bps) | -19.70*** | -5.26 | -20.51*** | -30.40*** | -7.94* | -33.43*** |
| | (1.37) | (4.12) | (1.61) | (5.09) | (4.01) | (4.59) |
| Observations | 74,092 | 72,013 | 51,560 | 50,392 | 48,293 | 38,706 |

*Notes.* Dependent variable: synthetic greenium (bps) at the bond–day level. All columns include bond and week fixed effects; standard errors clustered by country and week. Risk measures enter in units of 100 bps. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7. Between versus Within: Mundlak Decomposition

| Greenium measure | Component | $\hat{\theta}$ | SE | $t$ | Sample |
|---|---|---|---|---|---|
| Synthetic + CDS | Within | -18.79*** | 1.41 | -13.35 | All countries (N=35) |
| Synthetic + CDS | Between | +8.84*** | 1.74 | 5.08 | All countries (N=35) |
| Country-matched + CDS | Within | -18.63*** | 1.54 | -12.10 | All countries (N=36) |
| Country-matched + CDS | Between | +15.69*** | 4.50 | 3.49 | All countries (N=36) |
| Synthetic + CDS | Within | +3.17 | 4.08 | 0.78 | Excl. Egypt (N=34) |
| Synthetic + CDS | Between | -3.16 | 2.62 | -1.21 | Excl. Egypt (N=34) |

*Notes.* Mundlak decomposition of the CDS effect on the greenium with week fixed effects: the between component is the coefficient on the country mean of CDS, the within component on the deviation from that mean. Standard errors clustered by country. The excluding-Egypt coefficients are point estimates statistically insignificant under country-clustered standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

and is indistinguishable from zero, so we do not describe the cross-section as confirmed. The within coefficient in equation (7) is negative and large: $\hat{\theta} = -19.70$ per 100 basis points of CDS ($-30.40$ with the I-spread), robust in magnitude across first-stage constructions ($-19.7$ to $-22.9$) and confirmed by country-matching ($-19.40$); cell-matching points the same way at about a quarter of the magnitude, with fragile bootstrap inference (appendix Table B8), reflecting mechanical attrition of exactly the high-variance observations. When a sovereign's risk rises relative to its own mean, its greenium compresses.

Both margins rest on a single episode. Excluding Egypt moves $\hat{\theta}$ from $-19.70$ to $-5.26$ (insignificant) with CDS and from $-30.40$ to $-7.94$ with the I-spread; in the Mundlak decomposition both components lose significance and flip sign without Egypt. The bond-fixed-effects estimate excluding Egypt ($-5.26$) and the Mundlak within component excluding Egypt ($+3.17$) differ in sign because they are different specifications; the substantive point is the same, since both are statistically indistinguishable from zero. Figure 3 displays the leverage directly: excluding any country other than Egypt leaves $\hat{\theta}$ within $[-20.31, -19.64]$ under the synthetic construction and $[-20.29, -19.31]$ under country-matching, while excluding Egypt moves it to $-5.26$ and $-3.67$. The greenium–risk relation this section documents is therefore, in this sample, an Egypt phenomenon on both margins. Egypt is absent from the scarcity panel of Section IV, for lack of a market-size series, so the scarcity slope owes nothing to the observations that drive every estimate in this section, and the two results have opposite robustness profiles. Inference splits accordingly: the wild cluster bootstrap, which preserves pooled weighting, rejects sharply ($p = 0.0003$), while Ibragimov–Müller ($p = 0.128$) and the jackknife ($p = 0.171$), which weight countries equally, do not. This is exactly the divergence expected when the signal is concentrated in one country. The slope placebo (pseudo $p = 0.045$) indicates the compression is not manufactured by the construction, with the caveat that placebo CDS slopes have fat tails.

### *B. Stress Episodes*

Because risk levels are not comparable across sovereigns, the informative comparison is within a country's own risk run-up. We estimate country-by-country slopes on each sovereign's rising-CDS window, with bond fixed effects, week-clustered standard errors, and both greenium constructions. Table 8 and Figure 4 report the results.

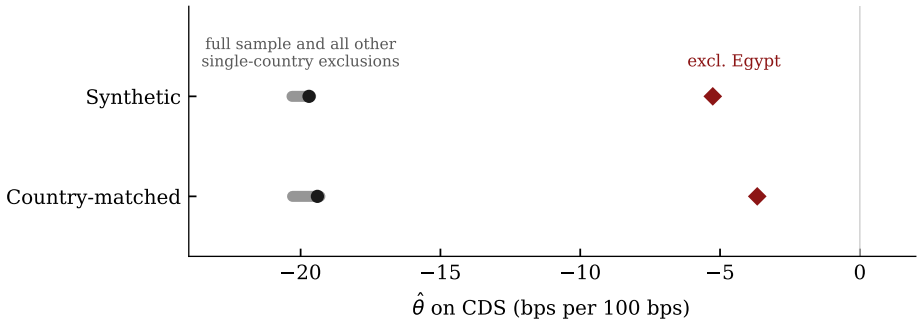


FIGURE 3. Leave-one-out sensitivity of the within-country risk slope. Grey bands span the estimates from excluding each country other than Egypt, one at a time; circles mark the full-sample estimate and red diamonds the estimate excluding Egypt, under the synthetic and country-matched constructions.

TABLE 8. Greenium Slopes in Sovereign Risk Run-Up Episodes

| | CDS rise | Synthetic | | Direct (matched pairs) | | |
|---|---|---|---|---|---|---|
| Country | (bps) | $\hat{\theta}_c$ | $t$ | $\hat{\theta}_c$ | $t$ | Classification |
| EGY | 1,554 | -23.94 | -9.7 | -23.97 | -9.6 | Robust compression |
| SRB | 252 | -16.70 | -8.7 | -11.62 | -11.0 | Robust compression |
| ROU | 98 | -21.97 | -5.9 | -2.18 | -0.6 | Suggestive (synthetic only) |
| PER | 94 | -7.31 | -1.6 | -18.81 | -2.4 | Suggestive (direct only) |
| MEX | 109 | -7.67 | -1.3 | -12.05 | -2.1 | Suggestive (direct only) |
| IDN | 93 | +11.14 | 1.0 | +6.89 | 1.7 | No compression |
| HUN | 212 | +19.33 | 11.1 | +11.66 | 5.7 | Widening |
| ITA | 103 | +8.71 | 2.3 | +15.11 | 5.6 | Widening |
| POL | 115 | +18.50 | 6.7 | +13.36 | 7.3 | Widening |
| LTU | 80 | +28.14 | 5.3 | +3.10 | 0.6 | Widening (synthetic) |

*Notes.* Within-country slope of the greenium on CDS (per 100 bps), estimated on each country's risk run-up window with bond fixed effects and standard errors clustered by week. Robust compression requires $t < -2$ under both greenium constructions. The Philippines and Malaysia are omitted because their green bonds start at the CDS peak; Chile, Côte d'Ivoire, and the Dominican Republic show no compression.

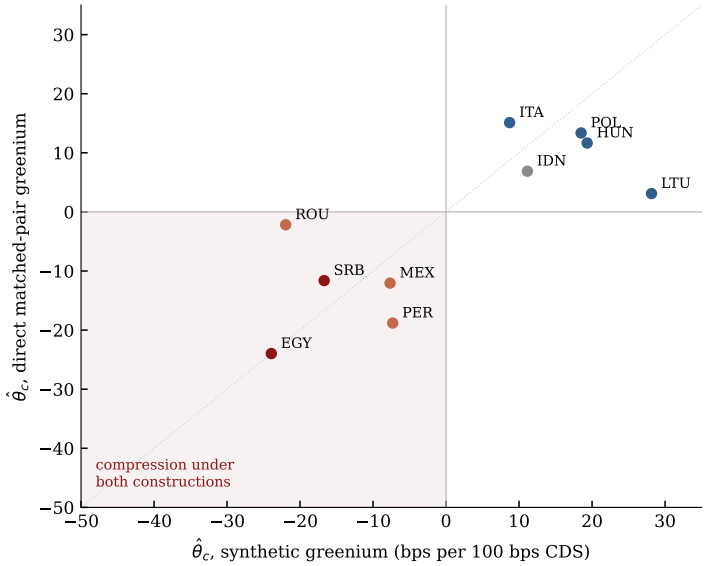


FIGURE 4. Within-country greenium slopes on CDS, full sample, under the synthetic (horizontal axis) and direct matched-pair (vertical axis) constructions. The shaded quadrant collects countries whose greenium compresses under both constructions. Slopes are basis points of greenium per 100 basis points of CDS. Each slope comes from a country-level regression with bond fixed effects and week-clustered standard errors.

There is no general within-country law. Over the full sample, 8 of 15 within slopes are negative, indistinguishable from coin flips (sign test $p = 1.00$; the full set is in appendix Table B7). The negative mean of $-4.28$ is generated by two outliers: it falls to $-0.10$ (median $+1.96$) once Egypt and the Philippines are excluded, the latter a country whose green bonds trade only from its CDS peak onward. What survives conservative treatment is a core of acute-stress episodes. Egypt, with the largest CDS run-up in the panel (1,554 basis points, peaking in October 2023, outside the 2022 global risk-off), compresses at about $-24$ basis points per 100 under both constructions. Serbia compresses under both ($-16.7$ synthetic, $-11.6$ direct), though its 2022 peak coincides with the global risk-off. Romania, Peru, and Mexico are suggestive under one construction each. Hungary, Italy, and Poland widen. The compressor core is robust to the greenium construction; individual-country conclusions in the moderate-stress group are not.

The scarcity framework, being static, predicts the between margin, which in this sample stands or falls with Egypt, and is silent on the within margin. The within compression is a new fact: in acute sovereign distress, the pricing distinction between green and conventional bonds of the same sovereign collapses, consistent with valuation becoming dominated by default probabilities, seniority and expected haircuts, and the exit of specialized green buyers, and with near-default pricing mechanics in the extreme observations. A state-contingent green premium, one that pays off in normal times and evaporates in distress, would rationalize both margins; formalizing it is beyond this paper. What the data reject is any reading of the risk prediction as an unconditional law that holds within countries at short horizons.

## VI. Issue Size and Market Depth

The framework's composition prediction operates at the project level: large, complex projects certify only where greeniums exceed the upper threshold $g^{\star}_{2,c}$. Sovereign green bonds, however, bundle many projects into one instrument, so the nominal size of an issue does not measure the size or complexity of any underlying project; a large French issue can fund many small efficiency upgrades, and a small Benin issue can fund one complex project. Issue size instead describes the scale of observed green bond placements.

With that distinction drawn, we treat the exercise as descriptive and auxiliary. Unlike the slope results in Section IV, it necessarily uses country-level greenium estimates in levels, whose sensitivity to counterfactual construction Section III.C has documented; the results should therefore not be read as a structural test based on precisely measured greenium levels. The exercise uses the broader country-level green-dummy construction and does not require the BIS market-size series used in Section IV, which expands coverage to 39 sovereign green issuers. Across those 39 sovereigns, continuous measures of relative green issue size are unrelated to the country greenium ($p \geq 0.63$). On the extensive margin, a probit for having at least one green issue larger than the country's median conventional issue loads *negatively* on the greenium ($-0.023$, $p = 0.008$): high-greenium sovereigns are less likely to have a large green issue. The pattern is driven by Egypt ($\hat{g}_c = +85.0$ basis points, no large issue) and Benin ($+78.5$, no large issue), the two highest-greenium sovereigns; countries without a large green issue have a mean greenium of $+10.4$ against $-1.5$ for the rest, but medians are indistinguishable (2.50

versus 2.68), so the pattern lives in the right tail rather than the whole distribution. Table 9 reports the specifications.

The issue-size results do not validate a literal translation of the project-composition prediction to sovereign issues. They nevertheless document that Egypt and Benin, the two highest-greenium sovereigns, have no large green issue. This pattern is consistent with thin green segments and limited certified supply, but does not establish that benchmark placements are infeasible or corroborate a separate market-size effect, which Section IV.B does not identify. The evidence remains auxiliary to Section IV; it cannot separate market size, liquidity, credit risk, and institutional capacity, and it depends on a handful of high-greenium cases. The project-level composition prediction remains untested and requires use-of-proceeds or project-characteristic data.

## VII. Concluding Remarks

We have asked when the sovereign greenium behaves as a scarcity price, measured the greenium at the bond level, and established by placebo which features of the measurement are reliable. The answer the data support is conditional. The greenium co-moves negatively with relative green supply, at about 2.5 basis points per percentage point, and this association persists in the pooled and emerging-market samples when relative supply is computed with a time-invariant, country-specific denominator. This is consistent with the framework's supply prediction, although the alternative ratios do not identify separate economic effects of green supply and market size. In the baseline specification the associa-

Table 9. Green Issue Size, Market Depth, and the Greenium

| *Panel A: issue-size regressions* Specification | Coef. | SE | $p$ | $N$ |
|---|---|---|---|---|
| MaxRel on greenium | 0.0025 | (0.0051) | 0.632 | 39 |
| MaxRel on greenium + log(TypicalSize) + AE | -0.0002 | (0.0058) | 0.970 | 39 |
| AvgRel on greenium + log(TypicalSize) + AE | 0.0008 | (0.0052) | 0.874 | 39 |
| Probit: Pr(HasLarge) on greenium | -0.0228*** | (0.0086) | 0.008 | 39 |
| Probit: Pr(HasLarge) on greenium + controls | -0.0224*** | (0.0083) | 0.007 | 39 |
| *Panel B: greenium by green-market depth* | | | | |
| | HasLarge = 0 | | HasLarge = 1 | |
| N countries | 20 | | 19 | |
| Mean greenium (bps) | +10.39 | | -1.50 | |
| Median greenium (bps) | +2.50 | | +2.68 | |
| SD greenium (bps) | +26.72 | | +13.49 | |

*Notes.* Cross-section of 39 sovereign green issuers. This exercise uses the broader country-level green-dummy construction and does not require the BIS market-size series used in Section IV. MaxRel (AvgRel) is the log ratio of the largest (average) green issue to the median conventional issue in the country's primary currency; HasLarge indicates that some green issue exceeds the conventional median. Country greeniums are estimated from first-stage regressions of z-spreads on a green dummy with week and currency fixed effects and tenor and bid–ask controls. Because this exercise uses country greenium levels, whose sensitivity to counterfactual construction is documented in Section III.C, the results are descriptive and auxiliary. Robust HC3 standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

tion is an emerging-market phenomenon and turns positive in advanced economies, a difference that survives country-level inference but not the imposition of common time shocks; whether the composition of supply and demand shocks differs across market types is an interpretation we do not document. The association also lives entirely in below-annual-frequency variation, indistinguishable from a flexible country trend. As Section IV.E explains, the negative supply gradient is consistent with scarcity pricing, but the low-frequency variation does not establish whether scarcity relief, changes in green demand, or broader market development account for the association. The relation between the greenium and sovereign risk runs across countries in the direction a viability-truncation mechanism implies, and within countries the greenium compresses in acute distress. But both margins rest almost entirely on Egypt and reverse without it. That concentration does not touch the supply results, which are estimated on a panel from which Egypt is absent. The highest-greenium sovereigns in the issue-size sample lack large green issues, a descriptive pattern consistent with thin green segments; the project-level composition prediction remains untested.

The findings inform market diagnosis and development rather than the effects of specific interventions. The framework isolates a scarcity component of the greenium, whereas the observed greenium may also reflect other pricing forces and measurement error. Moreover, although bond prices are observed daily, the empirical association is identified from slow-moving country-level variation rather than from responses to individual issues.

A first implication concerns measurement. The estimated level of the synthetic greenium is sensitive to the construction of the conventional counterfactual, while its slopes with respect to relative green supply and sovereign risk are more stable and are not reproduced by the placebo exercises. The greenium should therefore be measured using transparent and reproducible methods and interpreted jointly with issuance quantities and market characteristics. The evidence is more informative about how green-bond pricing covaries with market structure than about the precise financing advantage of a particular issue.

Second, the greenium is not an appropriate policy target. Under the supply-driven interpretation of the framework, expanding credible certified supply would reduce the scarcity premium. A declining greenium could therefore indicate an easing of the supply constraint rather than weaker market performance. The relevant outcome is the availability of credible financing for eligible investments, not the preservation of a yield differential.

Third, the results separate what the supply ratio does and does not establish. Holding market size fixed, variation in green supply reproduces the negative association with the greenium, so the association does not require fluctuations in the denominator. Holding green supply fixed, larger sovereign debt markets are associated with smaller greeniums, but Section IV.B does not identify this as a separate market-depth effect. Depending on the relevant constraint — thin certified supply or shallow, less developed markets — market development may involve strengthening the eligible expenditure or project pipeline and certification capacity, or improving repeated issuance, liquidity, benchmark size, and investor participation. The estimates do not identify which constraint is binding in each country.

The low-frequency nature of the association supports a programmatic rather than transaction-level interpretation. The baseline relationship becomes substantially smaller with country-year fixed effects and disappears with country-quarter fixed effects. It may therefore reflect persistent market devel-

opment and other slow-moving country characteristics rather than an immediate pricing response to individual green issues. From this perspective, public institutions and multilateral development banks could help address constraints on project preparation, certification, verification, and reporting while preserving the credibility of the label.

The findings also caution against a uniform interpretation across markets. Under the baseline time controls, the negative association is concentrated in emerging markets, while green supply and the greenium move positively together in advanced economies. This difference weakens under alternative time controls. Differences in the composition of supply and demand shocks are one possible explanation, but this interpretation remains untested. Policy interpretation must therefore account for market structure and issuance dynamics.

Finally, the sovereign-risk evidence illustrates the limits of green labeling. The compression of the greenium during increases in sovereign risk is driven largely by Egypt and is not a general within-country pattern. Where it occurs, it is consistent with the green attribute becoming secondary when default risk dominates pricing. A green label should therefore not be interpreted as protection against sovereign stress.

The evidence is associative throughout: issuance is endogenous to the greenium, lags provide no protection at the frequency where the variation lives, and mandated demand is unobserved, so the mandate-intensity prediction is untested. The natural next steps follow directly. A test of whether issuance responds differently to lagged greeniums in advanced and emerging economies would move the regime hypothesis from labeled conjecture toward evidence. Mandate data would allow the demand side of equation (3) to be brought to the data; plausibly exogenous variation in certification, such as the staggered adoption of green bond standards, would move the evidence toward identification; use-of-proceeds data would make the composition prediction testable; and a model with a state-contingent green premium would rationalize the within-country compression that the static framework leaves unexplained. Under the supply-driven interpretation of the framework, scarcity should erode as certified supply catches up with mandated demand, compressing the greenium.

## REFERENCES


Sakai Ando, Chenxu Fu, Francisco Roch, and Ursula Wiriadinata. Sovereign climate debt instruments: An overview of the green and catastrophe bond markets. IMF Staff Climate Note 2022/004, International Monetary Fund, 2022.

Sakai Ando, Chenxu Fu, Francisco Roch, and Ursula Wiriadinata. How large is the sovereign greenium? *Oxford Bulletin of Economics and Statistics*, 86(6):1472–1483, 2024.

Malcolm Baker, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler. The pricing and ownership of US green bonds. *Annual Review of Financial Economics*, 14:415–437, 2022.

Piotr M. Bolibok. Does firm size matter for ESG risk? Cross-sectional evidence from the banking industry. *Sustainability*, 16(2):679, 2024.

Juan Pablo Brichetti, Eduardo A. Cavallo, and Francisco Roch. The sovereign greenium: Global evidence and LAC perspectives. Technical note, Inter-American Development Bank, 2025.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427, 2008.

John Caramichael and Andreas C. Rapp. The green corporate bond issuance premium. International Finance Discussion Papers 1346, Board of Governors of the Federal Reserve System, 2022.

Stefania D'Amico, Johannes Klausmann, and N. Aaron Pancost. The benchmark greenium. SSRN Working Paper 4128109, 2023.

Raffaele Doronzo, Vittorio Siracusa, and Stefano Antonelli. Green bonds: The sovereign issuers' perspective. Markets, Infrastructures, Payment Systems Working Paper 3, Bank of Italy, 2021.

Alexander Dyck, Karl V. Lins, Lukas Roth, and Hannes F. Wagner. Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, 131(3): 693–714, 2019.

Torsten Ehlers and Frank Packer. Green bond finance and certification. *BIS Quarterly Review*, pages 89–104, September 2017.

Caroline Flammer. Corporate social responsibility and shareholder reaction: The environmental awareness of investors. *Academy of Management Journal*, 56(3):758–781, 2013.

Caroline Flammer. Corporate green bonds. *Journal of Financial Economics*, 142(2):499–516, 2021.

Gianfranco Gianfrate and Mattia Peri. The green advantage: Exploring the convenience of issuing green bonds. *Journal of Cleaner Production*, 219:127–135, 2019.

Robin Greenwood and Dimitri Vayanos. Bond supply and excess bond returns. *Review of Financial Studies*, 27(3):663–713, 2014.

Monika Grzegorczyk and Guntram B. Wolff. Greeniums in sovereign bond markets. Working Paper 17/2022, Bruegel, 2022.

Rustam Ibragimov and Ulrich K. Müller. $t$-statistic based correlation and heterogeneity robust inference. *Journal of Business and Economic Statistics*, 28(4):453–468, 2010.

Julia Kapraun, Carmelo Latino, Christopher Scheins, and Christian Schlag. (In)-Credibly green: Which bonds trade at a green bond premium? SSRN Working Paper 3347337, 2021.

Ralph S. J. Koijen and Motohiro Yogo. A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515, 2019.

David F. Larcker and Edward M. Watts. Where's the Greenium? *Journal of Accounting and Economics*, 69(2–3):101312, 2020.

Stefen MacAskill, Eduardo Roca, Benjamin Liu, Rodney A. Stewart, and Oz Sahin. Is there a green premium in the green bond market? Systematic literature review revealing premium determinants. *Journal of Cleaner Production*, 280:124491, 2021.

Yair Mundlak. On the pooling of time series and cross section data. *Econometrica*, 46(1):69–85, 1978.

L'uboš Pástor, Robert F. Stambaugh, and Lucian A. Taylor. Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571, 2021.

Ivan Sangiorgi and Lisa Schopohl. Explaining green bond issuance using survey evidence: Beyond the greenium. *The British Accounting Review*, 55(1):101156, 2023.

Dimitri Vayanos and Jean-Luc Vila. A preferred-habitat model of the term structure of interest rates. *Econometrica*, 89(1):77–112, 2021.

Olivier David Zerbib. The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking and Finance*, 98:39–60, 2019.

## APPENDIX A. CROSS-SECTIONAL PATTERNS

The pure cross-section of 29 country-average greeniums on average relative supply yields positive coefficients, opposite in sign to the within association, and quintile averages of the greenium are non-monotone in $\bar{s}$ (Table A1). Both are consistent with mandate–issuance simultaneity as a source of upward bias in the cross-section: where mandated demand is stronger, issuance is higher and, conditional on relative supply and other country characteristics, so is the greenium. This is a ceteris-paribus statement, not an unconditional ranking across countries; in particular, the advanced-economy markets in the sample can combine strong mandates with modest greeniums. Controlling for advanced-economy status raises the fit substantially, consistent with important cross-country heterogeneity that the 29-observation cross-section cannot isolate. We report it for completeness and rest nothing on it. Bond fixed effects absorb time-invariant country characteristics, but the within variation that identifies the slope is itself slow, so slowly moving country-level forces, including the mandate growth that confounds the cross-section, are not fully absorbed. The within margin remains more informative than 29 country means, but the two margins are less distinct than the high frequency of the bond panel suggests.

TABLE A1. Cross-Sectional Patterns of Scarcity and the Greenium

*Panel A: country cross-section ($N = 29$)*

| | (1) OLS | (2) WLS | (3) WLS+AE | (4) OLS | (5) WLS | (6) WLS+AE |
|---|---|---|---|---|---|---|
| $S_G/W$ (level) | 1.007* | 0.300 | 0.625** | | | |
| | (0.517) | (0.211) | (0.255) | | | |
| $\log(S_G/W)$ | | | | 2.762 | 1.943* | 2.105* |
| | | | | (1.702) | (1.064) | (1.030) |
| AE dummy | | | 4.391** | | | 2.796* |
| | | | (1.881) | | | (1.457) |
| Adj. $R^2$ | 0.067 | 0.000 | 0.105 | 0.143 | 0.091 | 0.132 |

*Panel B: greenium by quintile of relative green supply*

| Quintile ($\bar{s}$, %) | Mean $\hat{g}_c$ (bps) | SE | $N$ | Countries |
|---|---|---|---|---|
| Q1 (lowest) (0.26) | -9.17 | 7.27 | 6 | MEX, IND, IDN, KOR, CHN, LTU |
| Q2 (1.04) | -0.38 | 4.70 | 6 | ESP, AUS, SAU, MYS, PHL, POL |
| Q3 (2.40) | +1.07 | 2.08 | 5 | FRA, DEU, ITA, CHE, SWE |
| Q4 (4.55) | -0.98 | 3.55 | 6 | AUT, DNK, PER, THA, LVA, HUN |
| Q5 (highest) (9.03) | +3.51 | 1.23 | 6 | LUX, ISL, IRL, CHL, CYP, HKG |

*Notes.* Panel A regresses country-level greeniums on average relative green supply; robust HC1 standard errors. The positive coefficients are contrary to the within-country association and consistent with mandate–issuance simultaneity. Panel B reports the mean synthetic greenium by quintile of $\bar{s}$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

## APPENDIX B. ADDITIONAL TABLES

TABLE B1. Synthetic Greenium Panels

| Panel | Bond–day obs. | Green bonds | Countries | Mean (bps) | Median (bps) |
|---|---|---|---|---|---|
| Scarcity panel (BIS coverage) | 72,609 | 104 | 29 | 5.42 | 3.18 |
| Sovereign risk panel (CDS) | 74,092 | 112 | 35 | 7.33 | 3.41 |
| Sovereign risk panel (IMF I-Spread) | 50,392 | 79 | 26 | 7.22 | 2.01 |

*Notes.* Synthetic greenium panels underlying the analysis, at the bond–day level. The scarcity panel is restricted to countries with BIS general-government debt-securities data (the denominator of relative green supply); Egypt, among others, is absent from it. The risk panels merge the synthetic greenium with sovereign CDS (5-year) and the IMF I-Spread. Median first-stage $R^2$ across the 37 countries with an estimable first stage is 0.836. Means and medians in this table are computed after the corresponding risk merge; Table B2 reports first-stage means on the pre-merge construction sample.

TABLE B2. First-Stage Sensitivity of Slopes and Levels

| *Panel A: scarcity slope* $\hat{\delta}$ | | | | | | |
|---|---|---|---|---|---|---|
| First stage | $\hat{\delta}$ | SE | $t$ | $p$ | Obs. | Mean $\hat{g}$ (bps) |
| Baseline | -2.45 | 1.03 | -2.38 | 0.0244 | 72,609 | 5.42 |
| No bid-ask control | -2.21 | 1.01 | -2.19 | 0.0371 | 72,609 | 4.99 |
| Quadratic tenor | -2.23 | 0.96 | -2.33 | 0.0273 | 72,609 | 3.17 |
| Tenor spline | -1.96 | 0.90 | -2.17 | 0.0389 | 72,609 | 2.96 |
| Same currency | -1.64 | 0.50 | -3.28 | 0.0028 | 67,566 | 2.59 |
| *Panel B: sovereign-risk slope* $\hat{\theta}$ | | | | | | |
| First stage | Mean $\hat{g}$ | $\hat{\theta}$ (CDS) | SE | $\hat{\theta}$ (I-Spread) | SE | |
| Baseline (tenor + bid-ask; week [+ currency] FE) | 7.19 | -19.70 | 1.37 | -30.40 | 5.09 | |
| Same currency (tenor + bid-ask) | 4.88 | -19.65 | 1.55 | -29.56 | 6.19 | |
| Same currency, no bid-ask | 5.08 | -19.85 | 1.51 | -29.80 | 6.00 | |
| Same currency, quadratic tenor | 1.60 | -21.73 | 1.37 | -31.28 | 6.47 | |
| Same currency, maturity spline | -0.06 | -22.88 | 1.42 | -32.12 | 6.60 | |

*Notes.* Each row re-estimates the second stage varying only the first stage that constructs the synthetic greenium. Panel A: two-way (country×week) standard errors. Panel B: standard errors clustered by country and week. The mean synthetic greenium is highly sensitive to the first stage while both slopes are stable.

TABLE B3. Nonlinearity: Spline of the Scarcity Slope

| Segment | Range of $s$ (%) | $\hat{\delta}$ | SE | $p$ | Obs. | Countries |
|---|---|---|---|---|---|---|
| Low (< P20) | 0 - 1.1 | +10.59 | (9.07) | 0.253 | 14,385 | 25 |
| Middle (P20 - P80) | 1.1 - 10.5 | -4.68** | (2.00) | 0.027 | 44,054 | 22 |
| High (> P80) | 10.5 - 20.6 | -1.08*** | (0.35) | 0.005 | 14,170 | 4 |

*Notes.* Three-segment linear spline in $s$ with knots at the P20 and P80 of the bond–day distribution, bond and week fixed effects, two-way (country×week) standard errors. The low segment is dominated by countries with minimal green supply (China, Korea); the high segment by four countries. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

TABLE B4. Joint Log Decomposition on the Synthetic Panel (Collinear)

| Group | N | $\log S_{G,c}$ | $\log W_c$ |
|---|---|---|---|
| Pooled | 72,490 | -7.51 (6.15; $p$ 0.23) | +7.23 (11.27; $p$ 0.53) |
| EMDE | 49,747 | -8.36 (8.31; $p$ 0.33) | +8.52 (19.69; $p$ 0.67) |
| AE | 22,743 | -2.55 (4.58; $p$ 0.59) | -6.17 (6.59; $p$ 0.37) |

*Notes.* Joint decomposition $\hat{g}_{i,t} = \beta_S \log S_{G,c,t} + \beta_W \log W_{c,t} + \alpha_i + \gamma_t$ on the synthetic panel, two-way standard errors. No component is significant in any group; standard errors exceed the coefficients. This is collinearity, not evidence, and motivates the counterfactual-ratio design of Table 3.

TABLE B5. Numerator versus Denominator of Relative Green Supply

| | (D0) Ratio | (D1) Both | (D2) Only $S_G$ | (D3) Only $W$ |
|---|---|---|---|---|
| Green $\times \log(S_G/W)$ | -1.01 | | | |
| | (0.14) | | | |
| | [0.88] | | | |
| Green $\times \log S_{G,c}$ | | -0.73 | +0.15 | |
| | | (0.15) | (0.09) | |
| | | [0.68] | [1.16] | |
| Green $\times \log W$ | | +1.43 | | +0.93 |
| | | (0.14) | | (0.08) |
| | | [1.73] | | [1.73] |
| Observations | 454,092 | 454,092 | 454,092 | 454,092 |
| Wald $\gamma_1 + \gamma_2 = 0$, $p$ (week / two-way) | | 0.000 / 0.690 | | |

*Notes.* Dependent variable: z-spread; bond-level panel with country, week, and currency fixed effects and tenor and bid–ask controls, restricted to $S_G > 0$ and $W > 0$ (30 countries). D2 and D3 include the other variable as a control without the Green interaction. Standard errors clustered by year–week in parentheses; two-way in brackets. Stars follow two-way clustering; no interaction coefficient is significant under it. The joint specification (D1) suffers severe multicollinearity (VIF $\approx 6.3$). $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.10$.

TABLE B6. Green-Dummy Panel: Z-Spreads and Relative Green Supply

| | (1) Base | (2) Levels | (3) Log | (4) EMDE | (5) AE | (6) Lag 4w | (7) Lag 13w |
|---|---|---|---|---|---|---|---|
| Green | -3.01*** | -3.65*** | -2.07*** | -3.21*** | -3.22*** | -3.72*** | -3.69*** |
| | (0.15) | (0.25) | (0.23) | (0.31) | (0.64) | (0.25) | (0.23) |
| Green × supply term | | 0.15 | -0.73 | 0.22 | 0.26 | 0.15 | 0.10 |
| | | (0.04) | (0.16) | (0.04) | (0.16) | (0.04) | (0.03) |
| Observations | 455,083 | 455,083 | 455,083 | 259,990 | 195,093 | 449,976 | 437,196 |
| $R^2$ | 0.847 | 0.847 | 0.848 | 0.842 | 0.862 | 0.848 | 0.849 |
| Interaction $p$, two-way | | 0.637 | 0.340 | 0.403 | 0.701 | 0.637 | 0.739 |
| Interaction $p$, wild bootstrap | | 0.641 | 0.322 | 0.123 | 0.646 | 0.644 | 0.756 |

*Notes.* Dependent variable: z-spread at the bond–day level, with country, week, and currency fixed effects and tenor and bid–ask controls. The supply term is $s$ in levels (columns 2, 4–7, with lags in 6–7) and $\log s$ in column 3. Standard errors clustered by year–week in parentheses. Stars on the Green level follow year–week clustering; stars on interaction terms follow two-way clustering, under which, as under the wild cluster bootstrap, none retains conventional significance. $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.10$.

TABLE B7. Full-Sample Within Slopes of the Greenium on CDS

| Country | Green bonds | $\hat{\theta}_c$ (full sample) |
|---|---|---|
| EGY | 2 | -19.23 |
| SRB | 2 | -21.49 |
| ROU | 2 | -21.04 |
| PER | 2 | -14.43 |
| MEX | 5 | -7.45 |
| IDN | 6 | -24.88 |
| PHL | 3 | -43.64 |
| HUN | 8 | -4.19 |
| CHL | 12 | +1.97 |
| ITA | 5 | +4.89 |
| POL | 5 | +23.71 |
| CIV | 1 | +1.96 |
| DOM | 1 | +15.43 |
| MYS | 1 | +12.92 |
| LTU | 1 | +31.28 |
| Mean | | $-4.28$ |
| Mean, excl. EGY & PHL | | $-0.10$ |
| Median | | $+1.96$ |

*Notes.* Within-country slope of the synthetic greenium on CDS (per 100 bps) over each country's full sample, with bond fixed effects; this is the set of 15 slopes behind the sign test in Section V.B (8 of 15 negative, $p = 1.00$). The negative mean is driven by Egypt and the Philippines; excluding both, the mean is essentially zero and the median positive. Slopes for countries with one or two green bonds rest on thin within variation. Estimated country by country with bond fixed effects and standard errors clustered by week.

TABLE B8. Alternative Greenium Constructions: Sovereign Risk Slope

| Matching | Risk | $\hat{\theta}$ | SE | $t$ | Bonds | Countries | Reading |
|---|---|---|---|---|---|---|---|
| Synthetic | CDS | -19.70 | 1.37 | -14.34 | 112 | 35 | Preferred |
| Synthetic | I-Spread | -30.40 | 5.09 | -5.98 | 79 | 26 | Preferred |
| Country | CDS | -19.40 | 1.49 | -13.00 | 113 | 36 | Confirms synthetic |
| Country | I-Spread | -28.52 | 6.46 | -4.41 | 81 | 28 | Confirms magnitude |
| Cell | CDS | -4.97 | 0.66 | -7.51 | 103 | 35 | Directional; fragile bootstrap |
| Cell | I-Spread | -7.02 | 2.48 | -2.83 | 70 | 27 | Directional; fragile bootstrap |

*Notes.* All rows use bond and week fixed effects with standard errors clustered by country and week. Cell matching compares each green bond to conventional bonds in the same country×currency×maturity bucket and week; country matching to the average of all conventional bonds of the country in the week. Panel attrition in the cell estimator is mechanical (Hong Kong, Chile, and Poland lose 74, 20, and 27 percent of their synthetic-panel observations).

TABLE B9. Relabeling Placebo: Levels versus Slopes

| Object | Real | Placebo mean | Placebo median | Placebo SD | Pr(\|pbo\| $\geq$ \|real\|) | Percentile of real |
|---|---|---|---|---|---|---|
| Level (bps) | 7.19 | +1.60 | -0.10 | | | |
| Scarcity slope | -2.45 | -0.13 | -0.04 | 0.97 | 0.015 | 1 |
| CDS slope | -19.70 | +0.43 | -0.16 | 9.94 | 0.045 | 2 |

*Notes.* Conservative relabeling placebo, $B = 200$ iterations: in each country at most one third of conventional bonds are relabeled as false greens, removed from the brown-only first stage, and assigned a placebo synthetic greenium. The level row reports the placebo mean of $+1.60$ bps against the real panel mean; its 95% CI is $[1.58, 1.61]$ for the pooled mean and $[1.28, 1.91]$ across per-iteration means (the interval quoted in the text), two different objects. The slope rows re-estimate the second-stage regressions on the placebo panels; the construction reproduces roughly 22 percent of the level but essentially none of the slopes.

TABLE B10. First-Stage Fit by Specification

| First stage | Countries | Models | Min | P10 | P25 | Median | P75 | P90 | Max |
|---|---|---|---|---|---|---|---|---|---|
| Baseline | 37 | 37 | 0.121 | 0.654 | 0.767 | 0.836 | 0.928 | 0.974 | 0.989 |
| No bid-ask | 37 | 37 | 0.119 | 0.628 | 0.732 | 0.835 | 0.913 | 0.966 | 0.988 |
| Quadratic tenor | 37 | 37 | 0.124 | 0.660 | 0.771 | 0.903 | 0.947 | 0.977 | 0.989 |
| Tenor spline | 37 | 37 | 0.125 | 0.669 | 0.821 | 0.909 | 0.947 | 0.978 | 0.991 |
| Same currency | 37 | 45 | 0.121 | 0.630 | 0.772 | 0.860 | 0.928 | 0.963 | 0.989 |

*Notes.* Distribution of first-stage $R^2$ (brown-only regressions) across estimated models. Belgium, Uruguay, and Benin have at most one conventional bond and no estimable first stage. Only China (0.12) falls below 0.50; low first-stage fit adds noise to the synthetic greenium and attenuates second-stage slopes toward zero rather than biasing them.