Università Commerciale Luigi Bocconi
PhD School

PhD program in: Economics and Finance

Cycle: 36

Disciplinary field: SECS-P/09

Essays on Contagion, Climate, and Green Finance

Advisor: **Mariano Massimiliano Croce**

Co-advisor: **Riccardo Colacito**

PhD Thesis by

María José Arteaga Garavito

ID number: 3134348

Year 2026

Acknowledgements

I am deeply grateful to the many people who supported me throughout this process.

Most importantly, I am profoundly indebted to my advisor and co-author, Max Croce, for his mentorship and infinite encouragement. His high standards and creative approach to research continue to inspire me. I am also deeply grateful to Riccardo Colacito and Nicola Borri, whose support and guidance throughout the PhD have been invaluable.

I am especially thankful to Max Croce, Paolo Farroni and Isabella Wolfskeil, with whom I co-authored the chapter on pandemic risk, the first chapter of this thesis and the project that marked my first real dive into research during the PhD. Their rigor, reliability, and generosity made that early experience not only formative but also incredibly rewarding. I am also grateful to Max Croce, Riccardo Colacito and Biao Yang for their mentorship and collaboration on our work on climate finance. Their insights and dedication were essential to developing the second chapter and to my growth as a researcher. Collaborating with all of them has been a privilege.

I would also like to thank the Department of Finance at Bocconi University for its support throughout the program, and especially Elena Carletti for the guidance, candid feedback, and research opportunities she offered at important moments of my PhD. I continue to admire the example she sets.

I am endlessly grateful to my parents and my sisters Mari, Pau and Chayis, and my brother-in-law Robin, whose support has never wavered across borders and years, and to my friends, for keeping me grounded.

Finally, to my Ricardos: my husband and my son, thank you for being the brightest lights in my life. You give me strength, joy, and purpose. This journey would not have been possible without your love and patience.

Thesis abstract

This thesis studies how financial markets process risk, information, and institutional responses to systemic challenges. It brings together three essays that examine distinct mechanisms through which capital is allocated in the face of uncertainty. These include sudden health crises, persistent climate risks, and the emergence of new public financial institutions that aim to steer private investment. While the first two chapters focus on market responses to informational shocks, the third chapter investigates a structural evolution in financial intermediation.

The first chapter, **When the Markets Get C.O.V.I.D.**, analyzes how global financial markets respond to the emergence and diffusion of pandemic-related news. Using high-frequency data on over 15,000 COVID-19 medical announcements and real-time Twitter-based news diffusion across countries, the paper documents significant asset price reactions to contagion information. It shows that both equity and bond markets price contagion risk, and that exposure to pandemic news can explain heterogeneity in returns across countries and sectors. A high-frequency no-arbitrage model with time-varying betas formalizes the risk-premium implications of these findings.

The second chapter, **International Climate News**, turns to climate-related information as a priced risk factor in international financial markets. It introduces novel country-level indices of climate attention based on more than 23 million newspaper tweets collected across 25 countries. The empirical analysis shows that countries with greater exposure to global climate news experience currency appreciation, capital inflows, and persistent underperformance of high-emission ("brown") stocks. A dynamic international model incorporating priced climate news shocks and heterogeneous country exposure explains these results by linking climate attention to exchange rate movements and capital reallocations.

The third chapter, **The Rise of Green Banks**, focuses on institutional design and financial behavior in the domestic climate finance space. Green Banks are nonprofit or quasi-public financial institutions that mobilize private capital for clean energy investments by offering co-financing, guarantees, and concessional loans. Using a novel hand-collected dataset of project-level activity across U.S. Green Banks, the chapter provides the first empirical study of how these institutions deploy capital, interact with traditional lenders, and target disadvantaged communities. It also introduces a theoretical model in which a Green Bank faces leverage

constraints and coordination frictions when co-investing with private institutions. The model generates testable predictions about public-private capital structuring and the distributional consequences of climate investment.

Together, these three chapters reflect a broader research agenda focused on how capital markets react to risk and how institutional responses shape those reactions. From pricing pandemic and climate information to analyzing the financial behavior of emerging public intermediaries, the thesis provides new evidence on how risk, policy, and capital allocation interact in an increasingly uncertain and climate-constrained world.

Contents

Chapter 1

When the Markets Get CO.V.I.D.: COntagion, Viruses, and Information Diffusion

Co-authored with M.M. Croce, P. Farroni, and I. Wolfskeil.[1]

Abstract

We quantify the exposure of major financial markets to news shocks about global contagion risk while accounting for local epidemic conditions. For a wide cross section of countries, we construct a novel dataset comprising (i) announcements related to COVID19 and (ii) high-frequency data on epidemic news diffused through Twitter (Hassan et al. 2019's methodology). We provide novel empirical evidence about financial dynamics both around epidemic announcements and at daily/intra-daily frequencies. Analysis of contagion data and social media activity about COVID19 suggest that the market price of contagion risk is significant.

Keywords: Asset Prices, Pandemic Risk, Medical Announcements, Text Analysis.
JEL Codes: G12, G15, I10.

---

[1]Published in the *Journal of Financial Economics*, Volume 157, 2024, 103850. Full text and Internet Appendix available at ScienceDirect.

1.1 Introduction

COVID19 has manifested itself as a very aggressive and fast epidemic that—at the time of the first draft of this paper—brought major economic countries to their knees.[2] Given the fast-increasing contagion curve of COVID19 and its global scale, this epidemic event is challenging common economic policy interventions and depressing the global value of our assets, i.e., the financial wealth of millions of households all over the world.

Given that severe virus-related crises are expected to become more frequent, we find it relevant to use COVID-19-related data to ask the following broad questions about financial market reactions to viral contagion risk. First, what is the average impact of medical announcements on financial returns? Equivalently, is the diffusion of this information enhancing wealth or adding risk? Second, what is the market price of news risk related to global contagion dynamics? Third, can local contagion conditions help us predict expected returns?

Last but not least, can we use social media activity to measure the production and diffusion of information about epidemic risk? This question is important for at least two reasons. First, fast epidemic outbreaks catch investors off guard; hence, real-time indexes based on social media news may function as a useful predictive tool. Second, estimating multidimensional models requires many observations that we may gather by using high-frequency data instead of waiting for daily medical bulletins.

In this study, we address these questions by quantifying the exposure of major financial markets to news shocks about global contagion risk accounting for local epidemic conditions. For a broad cross-section of countries, we construct a novel data set comprising (i) medical announcements related to COVID-19; and (ii) high-frequency data on epidemic news diffused through Twitter (among others, see Hassan et al. (2019) and van Binsbergen et al. (2022)). Across several classes of financial assets and currencies, we provide novel empirical evidence about financial dynamics (i) around epidemic announcements, (ii) at a daily frequency, and (iii) at an intra-daily frequency. Formal estimations based on contagion data and social media activity about COVID-19 confirm that the market price of epidemic risk is very significant. Hence prudential policies that mitigate global contagion or local diffusion may be extremely valuable for financial wealth. More broadly, we offer a methodology for constructing a rich framework of information diffusion and information attention that future empiricists can adapt to examine future sources of global crises.

Interpretation of our findings. Before describing our findings in more detail, we must clarify how to interpret them. There is no doubt that the COVID pandemic was unprecedented in many dimensions (Baker et al. (2020)). Nevertheless, our high-frequency social-media–based

[2] To assess the severity of COVID19, see the March 11, 2020 WHO Director-General's opening remarks.

approach can also be informative when adapted to future pandemic events. Tracking pandemic-related sentiment and contagion dynamics can be fundamental to portfolio managers in future pandemic events. In this sense, our analysis is not simply a case study of a unique, rare event. Rather, it is very flexible as it shows how to gather a rich-and-reliable dataset for formal statistical tests even after a few weeks from the beginning of a 'brand new' pandemic. Given our high-frequency approach, we can track the full evolution of the pandemic across multiple different contagion waves. Furthermore, our methodology is broad as it allows financial economists to study many relevant dimensions of financial markets (see, for example, Hassan et al. (2020)).

In addition, after having collected more than 15,800 medical announcements across many countries, our results on the positive *average* appreciation of equity markets can be reasonably interpreted as a statement on *expected* appreciation. Our novel and sizeable dataset should minimize concerns about 'peso problems'.

Even though we aim to provide a flexible set of tools for future pandemic-related studies, we acknowledge that future crises may differ from the COVID one. In this case, future research should adopt our seminal approach as already done, for example, in the financial intermediation literature, to analyze rare-but-severe global financial crises.

Current results in detail. An important contribution of our work is the collection of a novel dataset on the COVID-19 pandemic that includes (i) an extensive set of official announcements on medical conditions (more than 16,000 announcements) and (ii) news diffused on Twitter in real-time by major newspapers (based on more than 823,000 tweets). We identify major newspapers for a large cross-section of countries in the spirit of Baker et al. (2016). We do not analyze articles; instead, we track news published on Twitter in real-time to produce high-frequency data when needed.

More specifically, we track tweets posted by major newspapers with keywords such as 'coronavirus' and 'covid19'. For each newspaper, we use the location of its headquarters to identify its specific time zone. As a result, we gather thousands of tweets for a large cross-section of countries that we can aggregate at different frequencies and across regions.

Given this data set, we document several important facts about news diffusion. First, both Twitter-based news diffusion (measured by the number of tweets) and attention (measured by the number of retweets) spike upon contagion-related announcements. Second and more broadly, the diffusion of information increases substantially in each country in our data set as soon as that country goes into an epidemic state.[3] Third, our measured increase in information diffusion is particularly pronounced precisely during the hours in which financial markets are

[3]We identify the beginning of the epidemic state with the day on which the number of confirmed COVID-19 cases becomes greater than or equal to 100.

open. All of these empirical facts suggest that tracking Twitter-diffused news is a reliable way to capture the information set of investors at a high frequency.

Turning our attention to financial dynamics, we look at equity returns around announcements, that is, in a ± 60 minute window. We find that cumulative equity returns jump upward in the post-announcement time window. This result is robust across several different specifications. In addition, we conduct the same analysis by looking at the government bond market and find a small decline in advanced economies (henceforth AEs) and no adjustment in emerging economies (henceforth EEs). We show that these results are consistent with a simple model in which the (i) demand of assets (see, for example, Koijen and Yogo (2019)) is driven by agents who care about the timing of resolution of cash-flow uncertainty; and (ii) the supply of bonds is less upward sloping than the supply of equities. Our high-frequency result is also consistent with the results documented by Gormsen and Koijen (2020) looking at dividend futures.

According to an LDA model applied to our tweets (in the spirit of Bybee, Kelly, Manela, and Xiu (2020b)), cases are one of the main drivers of the topics that received attention during the pandemic. Accordingly, in the last step of our analysis, we group countries into three portfolios daily according to their relative number of COVID-19 cases. We do this separately for AEs and EEs. The H (L) portfolio comprises the equity returns of the top (bottom) countries in terms of COVID-19 contagion cases. We then estimate a no-arbitrage based model in which we allow for time-varying betas ($\beta_{i,t}$) with respect to global contagion risk. Specifically, we allow equity returns to respond to global viral contagion news according to each portfolio's relative share of official COVID-19 cases. Global contagion risk is measured either by innovations in the growth rate of global COVID-19 contagion cases or by innovations in the tone of our COVID-19-related tweets.

This model can capture many of the features of equity returns that we document in our descriptive analysis. First, this model captures predictability through contagion-based time-varying betas. Second, this specification has the potential to capture higher negative skewness for countries that go through more severe contagion paths.[4] Third, this model accounts for heterogeneous exposure to global contagion news, enabling us to identify the market price of risk of this global contagion component.

Across all of our specifications, the market price of contagion risk is both statistically significant and extremely high. Equities are more exposed to risk than bonds. Both within AEs and EEs, heterogeneous exposure to contagion risk is substantial, and as a result, an equity-based HML-COVID strategy bears a high risk premium. An HML-COVID strategy that goes long in bonds of countries with a larger share of cases and short in those with a smaller share of cases, instead,

[4]Consider the case of portfolio H comprising countries receiving a sequence of relatively more severe contagion news. This portfolio will have greater exposure to adverse news ($|\beta_{H,t}|$ increases) as the relative contagion share of the portfolio grows. As the relative contagion share starts to flatten out and eventually decline, the sensitivity of this portfolio to good news is reduced ($|\beta_{H,t}|$ shrinks), meaning that returns will be less sensitive to positive news and hence the right tail of their distribution will not be very long.

provides an insurance premium. This suggests that bonds tend to become safer in countries exposed to heightened contagion risk. We find that this result is particularly sizable among EEs.

These results conform well with the data on weekly international investment flows. Countries with lower (higher) contagion levels are expected to experience equity inflows (outflows). Expected inflows are stronger in AEs than in EEs. In contrast, when looking at bonds, these findings are almost absent in AEs and reversed in EEs, meaning that in high-covid emerging economies, the flows going toward government bonds increase. This is consistent with the idea that bonds are perceived as safer assets in EEs.

In order to further exploit our rich framework, we consider a more granular cross section by looking at industry-level equity indices in Europe. This task is relevant for at least two reasons: (i) industries are well known to be difficult to price; and (ii) our industries comprise firms based in different countries and hence their riskiness is based on an interesting mix of country-level contagion conditions. Our estimation confirms that the market price of contagion risk is significant. In addition, we find relevant heterogeneity in risk exposure across industries.

In the last step of our analysis, we run intra-day regressions taking advantage of our high-frequency Twitter-based risk measure. We focus on European countries whose markets are open simultaneously, namely, ITA, ESP, UK, FRA, DEU, CHE, and SWE. Every day, we group them into three portfolios according to their relative number of COVID-19 cases measured in the previous 24 hours. The H (L) portfolio comprises the equity returns of the top-2 (bottom-2) countries for COVID-19 contagion cases.

Our novel high-frequency estimation confirms our main findings: contagion risk carries a significant market price of risk. Hence policies related to the prevention and containment of contagion could be valuable not only in terms of lives saved but also in terms of global financial wealth. These results also hold after controlling for the market and changes in equity volatility. Our results have been very stable over time and can be explored at `https://sites.google.com/view/when-markets-get-covid/`, a website that we use for the visualization of our data.

Related literature. Due to its relevance, the COVID-19 crisis has spurred a lot of contemporaneous research (Goldstein et al. (2021)). An important strand of the literature focuses on the measurement of both COVID-19-induced uncertainty and firm-level risk exposure by utilizing textual analysis and surveys (see, for example, Hassan et al. (2020) and Giglio et al. (2020)). We focus on high-frequency data, Twitter-based news diffusion, epidemic announcements, and country-level asset price dynamics.

Within the literature that studies news coverage and reaction to news, our manuscript is methodologically related to the work of, among others, van Binsbergen et al. (2022), Bianchi et al.

(2021), Hassan et al. (2019), Manela and Moreira (2017), Garmaise et al. (2021), Bybee et al. (2020a), Schmeling and Wagner (2023) and Engle et al. (2020).

Many studies look at the financial implications of COVID-19 (see, among others, Augustin et al. (2021), Bonaccolto et al. (2019), Bretscher et al. (2020a,b), Albuquerque et al. (2020), Ramelli and Wagner (2020), Pástor and Vorsatz (2020), Papanikolaou and Schmidt (2020), and Kaniel and Wang (2020)). In contrast to us, they do not focus on medical announcements and they do not assess the market price of viral contagion risk.

Several studies focus on firm-level implications (see, among other, Cororaton and Rosen (2020), Acharya and Steffen (2020) and Carletti et al. (2020)). Hartley and Rebucci (2020) and Sinagl (2020) look at monetary policy announcements and cash-flow risk, respectively. We differ in our attention to medical announcements; our social media-based measures of information diffusion and attention; and our high frequency analysis. Our work complements the evidence in Gormsen and Koijen (2020) and Gormsen et al. (2021) who extract relevant information about expectations and risk premia from derivatives.

Darmouni et al. (2023) studies the US corporate bond market fragility and assess policy interventions in times of COVID-19. The authors show that the weakness of mutual funds targeting corporate bonds has significantly exacerbated the drop in corporate bond prices. Our model is silent on the role played by risky bonds. We leave the analysis of the link between medical announcements and international corporate bond risk premia to future research.

1.2 Medical Announcements

In this section, we propose a simple model to think of asset demand around announcements. The model suggests that, on average, announcements should produce a reallocation from bonds to equities. As a result, equities should appreciate upon announcement, whereas bond prices should stay stable (decline) if their supply is flat (upward sloping). We test these predictions in our novel data set comprising thousands of COVID-19-related announcements across twenty-one countries. We then show our main results. Specifically, we document that: (i) equity markets, on average, appreciate upon announcements, and especially so in EEs; (ii) bond prices decline slightly in AEs, but stay stable in EEs; (iii) across both AEs and EEs, trade becomes more active upon medical announcements.

1.2.1 A Simple Model of Assets Demand and Announcements

Consider an agent with Epstein and Zin (1989) recursive preferences over two times in a period, $t = 0, 1$:

$$U_0 = \left[(1-\delta)C_0^{1-1/\psi} + \delta E_0\left[C_1^{1-\gamma}\right]^{\frac{1-1/\psi}{1-\gamma}}\right]^{\frac{1}{1-1/\psi}}.$$

For the sake of simplicity, consider the case in which the intertemporal elasticity of substitution is one, $\psi = 1$, and consumption at time 0 is equal to one, $C_0 = 1$. Without loss of generality, impose a subjective discount rate $\delta = 1$. If consumption is log-normal, the following applies:

$$U_0 = E_0[C_1^{1-\gamma}]^{\frac{1}{1-\gamma}} \approx E_0[C_1] - \frac{1}{2}(\gamma-1)V_0[C_1], \tag{1.1}$$

meaning that when $\gamma = 1$, the agent cares only about expected utility, $E_0[C_1]$, whereas when $\gamma > 1$ the agent dislikes uncertainty, $V_0[C_1]$. At time $t = 0$, the agent chooses how many bonds, B, and stocks, S, to buy taking as given prices along their supply curves. Think of B and S as the face value and the book value of bonds and stocks, respectively. The agent faces the following problem:

$$\begin{aligned} U_0(W_0) &= \max_{B,S} E_0[B+\theta S] - \frac{1}{2}(\gamma-1)V_0[B+\theta S], \\ W_0 &\geq p(B)B + p(S)S, \end{aligned} \tag{1.2}$$

where θ is a random variable that captures the riskiness of equity payouts. Since $V_0[B+\theta S] = S^2 V_0[\theta]$, the implied demand curves for bonds and equities are:

$$\begin{aligned} \frac{p(S)}{p(B)} &= E_0[\theta] - (\gamma-1)V_0[\theta]\cdot S \\ p(B) &= \frac{W_0 - Sp(S)}{B}. \end{aligned} \tag{1.3}$$

Assume that the supply of bonds is perfectly controlled by the central bank so that $p(B) = \overline{P}$. Without loss of generality, assume that $\overline{P} = 1$. Assume that the supply of equity is linear and upward sloping,

$$p^s(S) = a + \underbrace{b}_{>0} S, \tag{1.4}$$

where we impose $a < E_0[\theta]$. Under these conditions, at the equilibrium,

$$S = \frac{E_0[\theta] - a}{(\gamma-1)\sigma^2 + b} > 0. \tag{1.5}$$

We think of an announcement as an unbiased signal about θ that arrives at time $t \in (0,1)$ and that reduces the posterior uncertainty about equities, σ^2 (Ai and Bansal (2018)). The agent can freely revise her portfolio composition upon the arrival of the announcement. Since we can prove that,

$$\frac{\partial p(S)}{\partial \sigma^2} < 0 \quad \text{if} \quad \gamma > 1 \text{ and } \frac{\partial p(S)}{\partial \sigma^2} = 0 \quad \text{if} \quad \gamma = 1, \tag{1.6}$$

if the investor cares about the timing of information ($\gamma > 1$), *on average* announcements should be associated to equity appreciation as the investors shift their allocation toward equities. In what follows, we work with a novel dataset of covid-related announcements to test this hypothesis. In addition, we point out that if the supply curve of bonds is slightly upward sloping instead of being flat, we should expect to see both an average appreciation for equities and a depreciation for bonds, respectively.

1.2.2 Data Collection

We treat the release of each medical bulletin as an announcement. The same applies to travel limitations and lockdown policies related to COVID-19. We note that we have manually tracked these policy interventions on a daily basis and have constructed a novel dataset important to study real-time high frequency reactions of financial markets to epidemic risk.

At the beginning of our sample, we also witnessed important announcements related to both monetary and fiscal policy interventions. These announcements are not included in our study and, in addition, we also run our analysis excluding days with major monetary and fiscal policy announcements.[5] Hence we are confident that our results are not contaminated by other announcements.

Our data collection is very comprehensive, as documented in table 1.1, and it comprises more than 15,000 medical announcements. An example of a COVID-19-related announcement follows:

> **2020-03-14 15:35:00 CET; Vice President @Mike_Pence and members of the Coronavirus Task Force will hold a press briefing at 12:00 p.m. ET. Watch LIVE: `http://45.wh.gov/RtVRmD`**

We 'hand-collect' these announcements in several ways. First, for each country, we look for official press statements publicly available on the local Ministry of Health (MoH) webpage. Suppose the press statement does not have an official time stamp. In that case, we look for it on the official Twitter account of the MoH or other related government entities (for example, the Twitter account of the Prime Minister). If this second attempt fails, we also look at the Twitter accounts of major local newspapers and focus on news about medical reports. These

[5] As an example, here is an announcement related to a monetary policy intervention in response to COVID-19:
> **2020-03-18 15:35:00 CET; 23:05:00 CET; FT Breaking News; ECB to launch 750bn EUR bond-buying programme.**

Table 1.1: Summary Statistics for Announcements

Country	Total no. Announcements	Case Reports	Live Streamed	President/ Prime Minister	Regularly scheduled
AR	605	33%	64%	3%	90%
AU	678	78%	4%	1%	34%
BR	975	64%	26%	2%	54%
CA	791	58%	21%	18%	31%
CH	627	78%	9%	13%	73%
CL	896	59%	29%	3%	45%
CN	721	82%	3%	1%	75%
CN-HK	1,376	55%	2%	1%	100%
CO	1,006	58%	34%	8%	79%
DE	283	87%	1%	7%	61%
ES	570	83%	1%	17%	63%
FR	567	77%	16%	6%	84%
IN	759	89%	1%	1%	63%
IT	654	74%	17%	8%	81%
JA	332	59%	5%	5%	32%
KR	642	80%	1%	4%	72%
MX	1,803	10%	45%	21%	51%
NZ	457	61%	29%	7%	72%
UK	711	82%	11%	7%	68%
US	1,386	17%	54%	7%	33%
Total	15,839	64%	18%	7%	62%

Notes: This table shows summary statistics for COVID19-related announcements that we collect for a large cross section of countries. Our real-time data range from 01/01/2020 to 02/22/2022. For each country, we report the total number of announcements, the fraction of announcements that report the number of positive COVID cases, that are live streamed, and that are announced by the country's President or Prime Minister. In the last column, we report the fraction of announcements that took place at a daily regular time.

steps, which we repeat multiple times each week, are sufficient to identify the effective time of each announcement in our data set relevant for financial investors.

Our dataset comprises mostly regularly scheduled announcements, that is, daily releases at a pre-established well-known time. We also collected irregularly pre-scheduled announcements. As an example of an irregularly pre-scheduled announcement, in figure 1.1, we report our record of the first scheduled Coronavirus Task Force Press briefing. In contrast to the subsequent White House press meetings, this briefing occurred earlier, at 3:40 p.m. EST. Note that this meeting was announced to the public 2.5 hours before it took place, hence it must be considered a pre-scheduled announcement. The country-level fraction of regularly scheduled announcements is reported in the last column of Table 1.1. We run our tests with and without irregularly pre-scheduled announcements and show that our main results hold in both cases.

More broadly, our dataset comprises an extensive set of purely COVID-19-related announcements, and it enables researchers to easily identify each specific announcement. In table 1.1, we

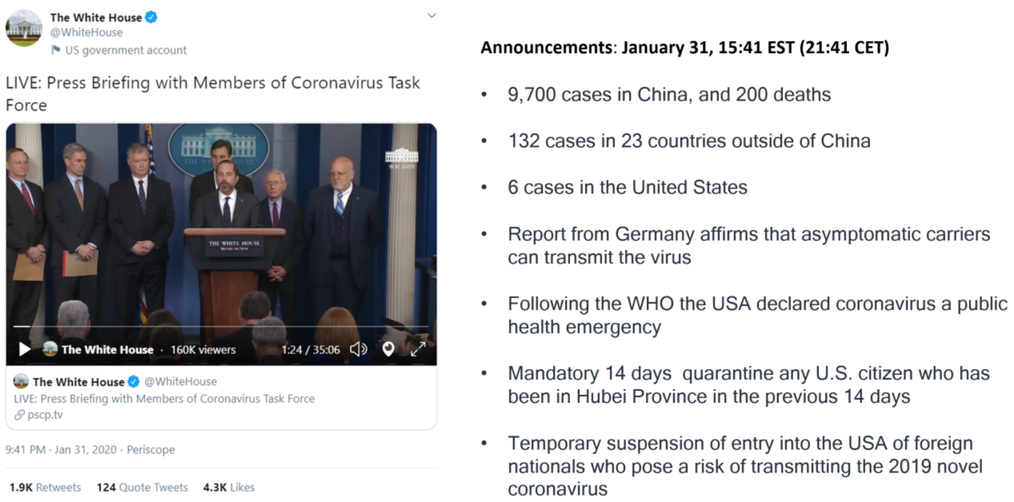


Figure 1.1: Announcement Time from Twitter.

Notes: This figure shows a tweet about one of the first COVID-related announcements in the US. The tweet time stamp enables us to identify the effective timing of the announcement. On the right hand side of this figure, we summarize the topics discussed during the briefing.

provide some interesting dimensions of our data set by detailing the share of announcements related to reports about Covid cases, live-streamed press conferences, and announcements by the President or the Prime Minister.

1.2.3 Announcements and Financial Markets

Pre- and post-epidemic samples. In what follows, we study the financial dynamics around medical announcement times. In order to isolate the dynamics related solely to medical announcements, we plot the differential behavior of our variable of interest with respect to normal times, i.e., pre-epidemic times. In each country, we define the beginning of the epidemic period as the day in which the country experienced an official number of contagion cases greater than or equal to 100. Given this threshold, China is the first country in our sample to go into the epidemic phase, whereas New Zealand is last. The appendix shows that our main results remain unchanged when we start the epidemic sample for all countries when China reached 100 cases, i.e., at a common date (see Appendix A).

The pre-epidemic sample starts for all countries on October 1st, 2019, so the pre-epidemic period comprises at least four months of data. This subsample is long enough to run meaningful comparisons with the post-pandemic subsample. Consider, for example, an announcement on a Friday at 3:40 p.m. EST. We compare the reaction of our financial variables around this announcement to their behavior at the same time in our pre-epidemic sample.

Pre- and post-announcement behavior. We run a high-frequency analysis around announcement times. In what follows, we estimate the following regression at the minute-level:

$$Z_t = (c_{pre} + c_{t>t^*}) + (\alpha_{pre} + \alpha_{t>t^*}) \cdot t + (\beta_{pre} + \beta_{t>t^*}) \cdot t^2, \quad t \in [t^* \pm K] \tag{1.7}$$

where t^* is the time of the announcement, K is equal to 60 minutes; and Z_t is the differential behavior of our variable of interest across the pre- and post-epidemic sample. By doing so, we remove the need of introducing country and day fixed effects. This specification is a quadratic function of time that includes dummy variables to account for post-announcement jumps in both the level ($c_{t>t^*}$) and the slope ($\alpha_{t>t^*}$, $\beta_{t>t^*}$). We test the null assumption that there is no difference post-announcement, $H_0 : c_{t>t^*} = \alpha_{t>t^*} = \beta_{t>t^*} = 0$, and if we fail to reject the null we depict the resulting smooth quadratic fit. Standard errors are always HAC-adjusted.

Financial data sources. All data are from Thomson Reuters and Bloomberg. Equity, bond, and currency data are obtained at the minute frequency and then aggregated at lower frequencies when necessary. For each country, we collect data on its major equity index and 10-year maturity treasury bond index. We measure the risk-free rate by focusing on the yield of 3-month government bills. Due to data availability, CDS data are collected at the daily frequency. All details about our data can be found in table A.3 (see Appendix A).

Equity markets. In figure 1.2, we show the average cumulative returns obtained from buying country-specific equities 60 minutes before a country-specific announcement and holding them for 120 minutes. Our results are averaged across both countries and announcements. Countries are divided into two groups, AEs and EEs, according to the IMF classification.[6]

The top panels show what happens when considering all countries and all announcements. Both in AEs and EEs, equity values appreciate substantially upon the announcement, consistent with our simple model presented in section 1.2.1. This appreciation is persistent, as it remains almost constant during the next hour in AEs and gets amplified in EEs. This observation suggests that the release of COVID-19-related news helps equities. Since we are considering a large number of announcements conveying both positive and negative news, we think of this jump in equity valuation as a measure of expected appreciation due to the reduction of uncertainty on epidemic risk (in our simple model, think of a reduction of the posterior variance, σ^2).[7]

One potential concern related to our results is that they may be driven by a peso problem.

[6] If a country-specific announcement happens when the exchange of the country is closed, we consider the 60 minutes prior to the closing time of the previous day and the first 60 minutes after the opening of the exchange in the next day.

[7] Lucca and Moench (2015) show a slow and persistent accumulation of positive returns before monetary policy announcements. This drift may be explained by information leakage. In our case, instead, the sudden increase in the cumulative returns at the announcement is consistent with no information leakage.

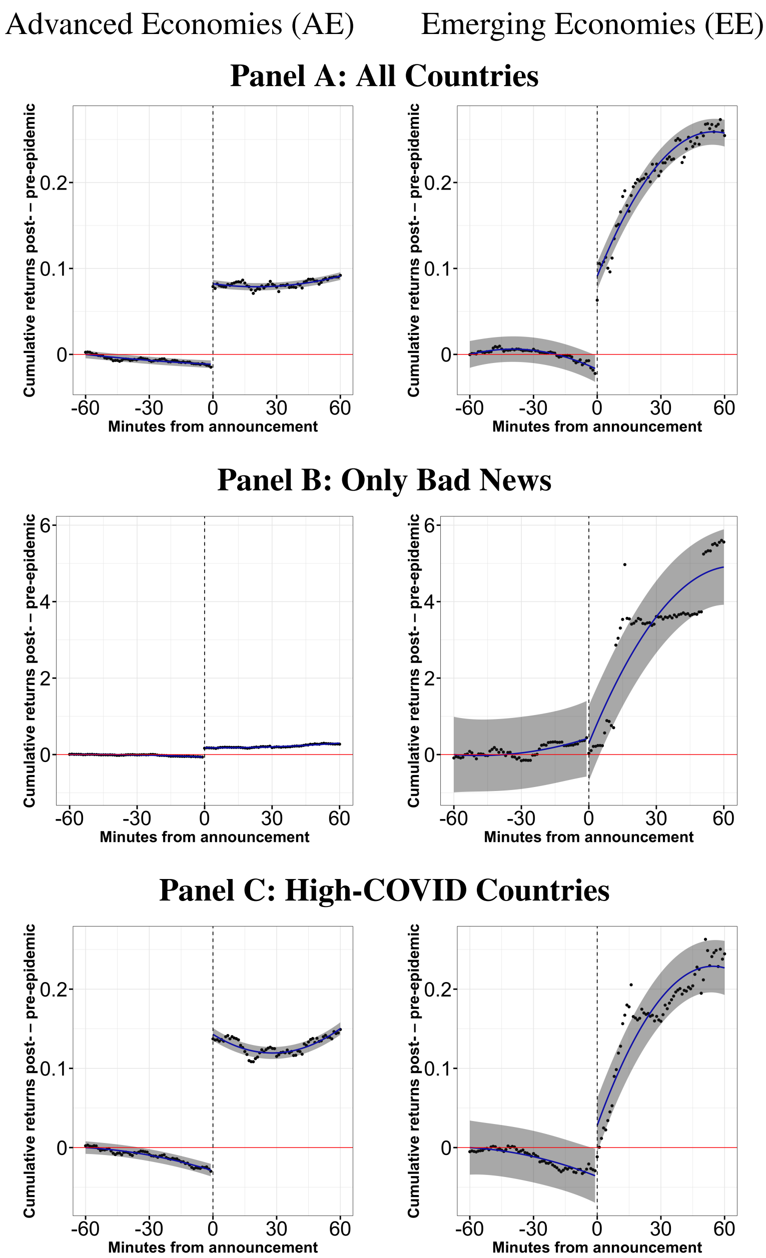


Figure 1.2: Equity Returns around Announcements

Notes: In each panel, dots denote the difference across subsamples of the cross-country-cross-announcement average cumulative returns obtained from buying equities 60 minutes before an announcement and holding them for 120 minutes. Panel a (c) comprises announcements from all countries (top-50% countries in terms of contagion cases) in each group. Panel b excludes announcements conveying good news. Returns are in log units and multiplied by 100. Solid line and shaded areas are based on the estimation of equation (1.7). Our sample starts on 10/01/2019 and ends on 02/22/2022.

In order to address this problem, we do two things. First, we work with a very large cross section of announcements. Second, we repeat our analysis by focusing only on announcements conveying bad news. We measure bad news as an unexpected increase in the growth rate of contagion cases on the day of the announcement. We explain in detail our construction of the news in the next section when we price them using the cross section of equity and bond returns. In figure 1.2, middle-left panel, we show that the same phenomenon is present to a similar extent when we focus on the subset of announcements associated with bad news within the group of AEs. Note that the scale for this panel is one order of magnitude greater than that in figure 1.2, top panel. Hence our average results for AEs are not driven by a few extreme realizations.

We further support this point by replicating our analysis on a subsample in which we remove days with news in either the top- or bottom-1% of our distribution (see figure A.1 in the Appendix). Our results are unchanged (see figure A.2 in the Appendix, panel a). In addition, our results are qualitatively unchanged if we focus on a subsample ending in December 2020, i.e., at the peak of the second COVID wave (figure A.2, panel b).

Turning our attention to EEs, we point out that in this case, the jump is one order of magnitude greater than under the case in which we consider all announcements. In EEs, it is clear that the average appreciation that we see in the hour after the announcements is not driven by the arrival of positive news.[8] In general, one could be concerned that our results are driven by an in-sample predominance of good news, resulting in more equity appreciations than depreciations upon announcement. In order to address this concern, we estimate the average of our news across countries and days. According to standard t-test, we cannot reject the null assumption that our news are on average zero. Equivalently, our novel and extensive data set comprises a similar number of good and bad news across a large cross section of countries and a long daily time series.

In figure 1.2 bottom-panel, we consider all of our announcements, but we limit our attention to countries above the median in total contagion cases. The scale in these panels is identical to that used in figure 1.2 top-panel. Not surprisingly, the smaller sample we use produces estimates surrounded by higher estimation uncertainty. Taking this into account, the value of the information disclosed during these announcements is higher among high-COVID AEs and remains almost unchanged among high-COVID EEs. More broadly, when we look at the entire cross section of our 21 countries, low-COVID countries appear to be less sensitive to contagion-risk news. This is consistent with the results of the no-arbitrage factor model that we estimate in the second part of our study.

The equity returns patterns that we document may also be consistent with models featuring behavioral attributes and micro-frictions. In order to provide more data to distinguish across theories, we also look at equity trade volume. In figure 1.3, we directly depict the difference in the log-growth of trade volume across normal and epidemic subsamples. We find that both in AEs and in EEs trade volume features no change before the announcements. Consistent with previous studies (see, among others, Han (2020)), trade volume increases right after the announcement. This upward adjustment is more pronounced in EEs. We interpret these results as strongly supporting the relevance of information arrival in markets where investors care about

[8] the first ten minutes after the announcement, negative news on average dominate the announcement premium. We depict our results for announcements associated to good news in figure A.2, panel c.

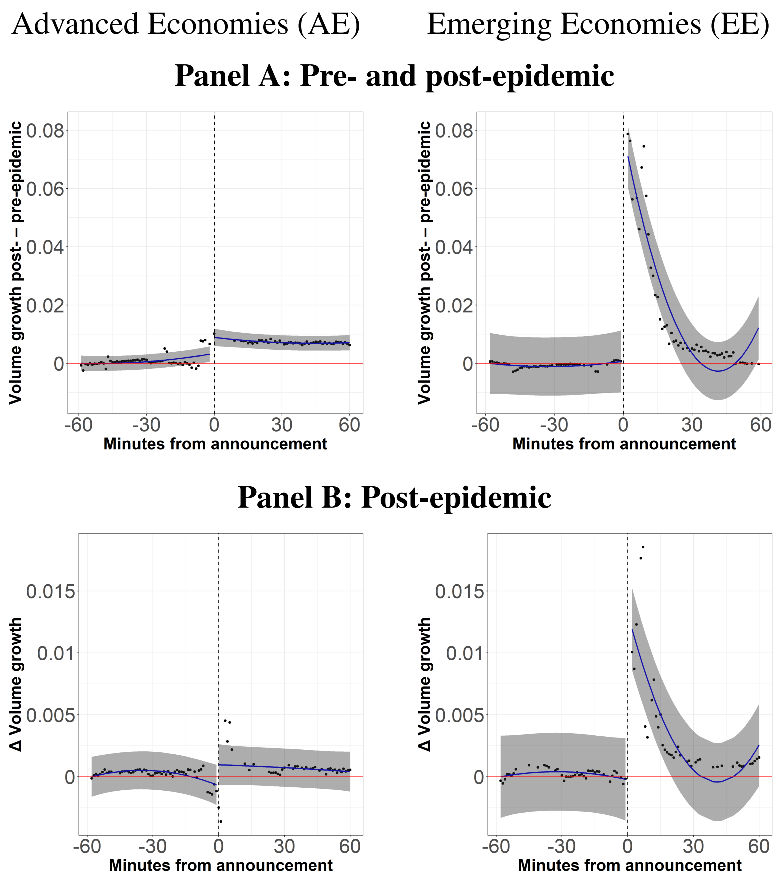


Figure 1.3: Equity Trade Volume around Announcements

Notes: The figure shows the average equity log-volume growth for all countries around announcement times. In the top panel we depict the difference across pre- and post-epidemic samples. In the bottom panel we plot the difference in volume growth for the post-epidemic sample between days when the absolute value of contagion news is above median and days when it is below median. In each country, the epidemic period starts when there are more than 100 cases of COVID19. Solid line and shaded areas are based on the estimation of equation (1.7). Our sample starts on 10/01/2019 and ends on 02/22/2022.

the timing of uncertainty resolution (see, for example, Ai and Bansal (2018), Ai et al. (2022a) and Ai et al. (2022b)).

In addition, these models predict that volume should be more pronounced when the announcements carry stronger news. In theory, this should apply to both positive and negative news. In order to examine this prediction, we look at the absolute value of the surprise in the number of cases across different days. In the bottom panels of figure 1.3, we show that volume increases relatively more on days with more extreme (unexpected) cases news. In the next part of this study, we focus on sovereign bonds and document that liquidity seems to increase in the bond markets as well.

Bond markets. Figure 1.4, top-panel, shows our results for bond returns. The construction of the depicted data is identical to that used for equities. In a ±60-minute window around the announcement, there is no significant adjustment in bond returns for EEs, consistent with

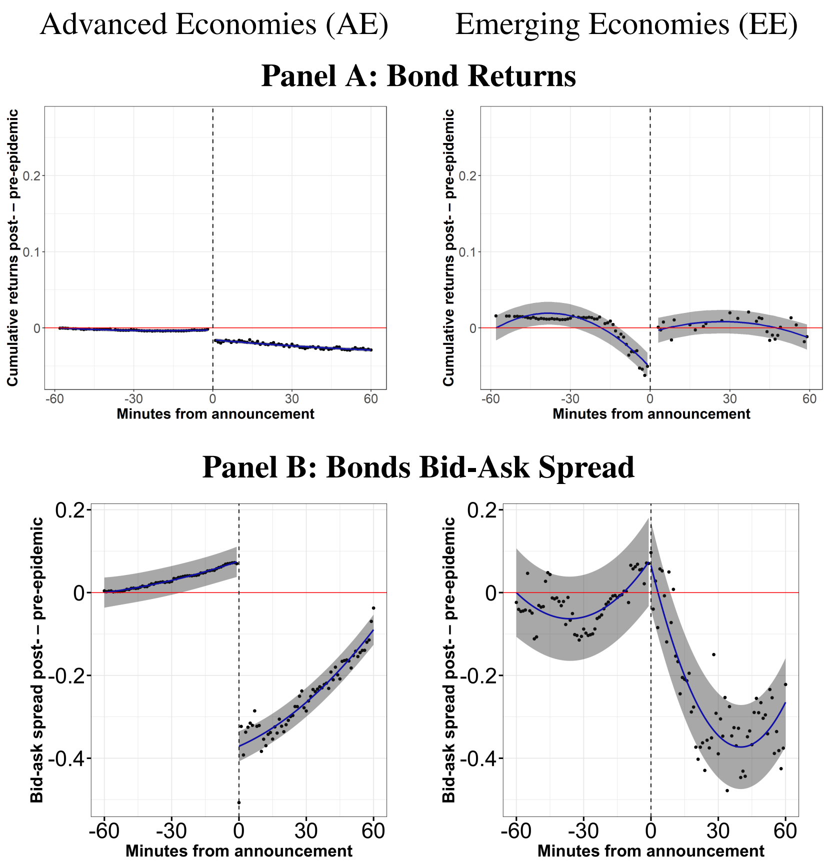


Figure 1.4: Sovereign Bonds around Announcements

Notes: In the top panels, dots denote the difference across subsamples of the cross-country-cross-announcement average cumulative returns obtained from buying 10-year sovereign bonds 60 minutes before an announcement and holding them for 120 minutes. In the bottom panels, dots refer to the difference across subsamples of the cross-country-cross-announcement average of the bid-ask spread of the bonds. Returns are in log units. All series are multiplied by 100. Solid line and shaded areas are based on the estimation of equation (1.7). Our sample starts on 10/01/2019 and ends on 02/22/2022.

our model in Section 1.2.1. In AEs, the adjustment is modest and negative, consistent with our simple portfolio allocation model if we assume that the supply curve of bonds is slightly upward sloping.

Through the lens of our model, these results suggest two important lessons. First, during the COVID-19 crisis, cash-flow uncertainty was an important determinant of the equity market. This high-frequency result is consistent with the results documented by Gormsen and Koijen (2020) looking at dividend futures. Second, given that their cumulative return is nearly zero across AEs and EEs, bonds are an important hedge against contagion risk announcements.

According to our model, announcements should prompt a trade away from bonds. Absent high-frequency data on bonds trading volume, we test this hypothesis by looking at their bid-ask spread which we interpret as a measure of inaction in the market. On average, after an announcement, there is an immediate decline in the bid-ask spread in AEs and a delayed one in EEs. This is consistent with the idea that covid announcements have been an important source

Table 1.2: CDS Spreads and Contagion News

	A.E.		E.E.	
Contagion cases - news	6.133***	7.735**	27.897***	27.290***
	(1.988)	(3.782)	(8.298)	(8.364)
Adj. R2	0.02%	4.64%	0.20%	14.24%
Adj. R2 w/o	0.02%	4.64%	0.20%	14.24%
Country FE	Yes	Yes	Yes	Yes
Week FE	No	Yes	No	Yes

Notes: this table reports the results of the following regression:

$$\Delta S_t^i = d_0^i + d_t^i \cdot D_t^{Week} + \beta^g \cdot news_{t-1} + \epsilon_t^i, \quad \forall i \in g$$

where ΔS_t^i refers to the daily change of the CDS spread in country i; g refers to either the group of Advanced Economies (AEs) or that of Emerging Economies (EEs); d_0^i is a country-level fixed effect and D_t^{Week} is a weekly time fixed effect. 'Contagion cases - news' refers to the innovation in the growth of the global number of contagion cases as measured in section 1.3. 'Adj. R2 w/o' refers to the adjusted R squred from the same regression in which we omit the contagion news. Standard Errors are clustered at the country-level. Our sample starts on 02/01/2020 and ends on 02/22/2022.

of trade.

An alternative explanation for the muted response of bonds is that they are subject to two offsetting forces. Specifically, flight to safety may promote bond appreciation, but sovereign default risk may increase and push bond prices downward simultaneously. In order to study the plausibility of this hypothesis, we collect daily country-level data on CDS spreads and link their daily variation to the daily news on contagion cases. We explain in detail how we measure news in the next section. Since different countries entered this crisis with different levels of fiscal capacity, exploring country-level heterogeneity is essential. For this reason, in our empirical analysis, we include both country-level fixed effects and week-level time fixed effects.

In table 1.2, we show that adverse contagion news tends to increase CDS spreads in a statistically significant way. This effect is three times stronger in EEs. Simultaneously, we document that this news produces a very modest increase in the adjusted R-squared of our regression, implying that for AEs, default concerns have been a second-order issue.

1.2.4 Additional results

This section summarizes a list of additional results reported in detail in Appendix A (available here). Since our benchmark analysis is based on several methodological choices, we show the role played by each of them in what follows. More broadly, we show that our results are robust to several methodological changes.

The role of country-specific epidemic dates. In our benchmark analysis, we have country-specific dates defining the beginning of the pandemic. As an alternative, we can pick a common date, namely the day on which China reported 100 official cases. Our results are unchanged when we adopt this strategy. As an example, see our results for equity in figure A.3.

Post-epidemic dynamics. In our analysis, we plot the differential behavior of equity markets around announcement times across our pre- and post-epidemic samples. We also report our results for equities in the post-epidemic sample (see figure A.4) and confirm that medical announcements, on average, are associated with an appreciation of equities. In addition, our model suggests that the realized jump in cumulative returns should decrease with bad news, i.e., it should be low (high) when the number of cases announced is above (below) expectations. In figure A.5 in the appendix, we depict the difference in cumulative returns across days in which countries experienced an unexpectedly high number of cases (bad news) and days in which they experienced an unexpectedly low number of cases (good news). The model suggests that this difference should be negative around the announcement time, and our figure confirms this result.

The role of domestic announcements. Recall that our cross section comprises 21 countries. We can think about the previous results about equity (bond) returns as the equal-weighted cumulative returns that an investor could obtain by trading ahead of each announcement across 21 sources of announcements (one per country) and in 21 equity (bond) markets, for a total of 21×21 possible trade combinations.

In order to disentangle the effects of local announcements on local markets, we also consider the average cumulative return of an investor that trades only in the domestic market ahead of domestic announcements. In figure A.6, we focus on the average cumulative returns across 21 trade strategies that involve neither foreign news nor foreign assets. Our data confirm that bonds have a muted response to announcements, whereas equities appreciate afterward.

Covid vs. macroeconomic announcements. In order to further isolate the role of medical announcements, we have created a dataset comprising the dates on which either inflation, industrial production, or GDP data are released in each country in our cross section. Our results continue to hold when we exclude these days from our dataset (see, for example, figure A.7). In other words, our announcement results are unrelated to other announcements already explored in the literature in terms of content (we focus on medical announcements) and timing (our results also hold on days when there are no major macroeconomic announcements).

Regularly pre-scheduled announcements. At high frequency, all of the announcements that we have utilized so far are pre-scheduled. This means that investors are aware that an announcement is going to be made within a few hours. Still, some of our announcements are not regularly scheduled, in contrast to macroeconomic announcements. In order to be consistent with previous studies, for each country, we identify regularly scheduled announcements by ensuring that (i) they are about case reports or live-streamed events and (ii) they are released at a recurrent country-specific time of the day. Our results also hold if we focus on regularly pre-scheduled announcements (see figure A.8). In addition, when focusing on regularly pre-scheduled announcements, our sample excludes almost entirely observations from the first six weeks of pandemic diffusion. Therefore, our results are not driven by early announcements of expansionary fiscal and monetary interventions or those discussing the spread of the pandemic in early-affected foreign countries like China and Italy. This subset of announcements primarily centers on the domestic number of cases.

Information Diffusion. Our novel social media-based data set enables us to measure the diffusion of information at a very high frequency. For each announcement in our data set, we compile all COVID-19-related tweets issued in a $\pm$60-minute window around announcement time by major newspapers in each country. In the next section, we provide a detailed description of our data collection procedure. For the sake of statistical power, we aggregate all of these tweets across all of our countries and call the resulting aggregate 'World'.

In the left panel of figure A.9, we show the per-country per-minute average cumulative number of tweets around announcement times during the post-epidemic sample. The right panel refers to retweets, that is, our measure of attention to the news. Both information diffusion and attention to the news increase significantly in the hour after announcements. In order to interpret the magnitude of the diffusion, we must clarify that we collect only tweets from major newspapers and their respective number of retweets. Because of API limitations, it is impossible for to us measure how many users read an individual tweet. Hence our figures represent a lower bound on the overall diffusion.

In addition, it is evident that most of the twitting activity takes place in the time window [-90 -30] ('preview' tweets about the announcement) and [0 +90] ('rehash' tweets). These results are reassuring as they provide additional support to our covid-related data set, as the time of our announcements nicely aligns with that of the news media cycle.

1.3 Contagion News

In this section, we attempt to price news about pandemic risk. We do it using two fundamental measures, namely, unexpected changes in the number of contagion cases and unexpected changes in the tone of the news about contagion. The first measure is based on an objective count of COVID-19 positive cases. Nevertheless, across different months or contagion waves, the same variation in the number of cases may be associated with different assessments of risk. For this reason, we find it necessary also to study a media-based measure of news tone.

Our analysis confirms that global epidemic news has a significant market price of risk. In April 2020, at the peak of the first COVID contagion wave in AE, daily equity risk premia may have increased by 28% in AEs and by 13% in EEs compared to the median risk premia in our sample. This disruption is comparable to that measured during the global financial crisis.[9]

1.3.1 Data Collection

Twitter-based news. In the spirit of Baker et al. (2016), we identify major newspapers for a large cross section of countries (see table A.1 in the appendix). In contrast to Baker et al. (2016), we do not analyze articles; rather, we track news published on Twitter in real-time to produce high-frequency data when needed. More specifically, we track the news related to the COVID-19 pandemic posted by major newspapers on Twitter. We do so by searching for keywords such as 'coronavirus' and 'covid19'. For each newspaper, we identify its headquarters location so we can determine its specific time zone.

In table 1.3, we report a summary of our social media–based dataset. It is very comprehensive and it features several dimensions that enable us to study both information production and diffusion. Specifically, our ability to track retweets and likes gives us a high-frequency measure of attention. Google searches are often used to measure attention (Da et al. (2011); Ramelli and Wagner (2020)), but to the best of our knowledge, they are not provided minute-by-minute, and they do not account for the timing of initial production of the news, an aspect that is very important when analyzing capital market reactions.

The time series behavior of our news indicator is depicted in figure 1.5. For each country, we also depict the beginning of the epidemic period, which we identify as the day the number of confirmed cases of COVID-19 exceeds 100. We note several interesting patterns. First, there is significant heterogeneity across countries in the timing of information diffusion. Across several countries, information diffusion becomes more intense after the beginning of the local epidemic period. We note that both the diffusion of news, that is, the number of tweets, and the attention to the news, that is, the number of retweets, increase rapidly after the beginning of the local epidemic period.

Figure 1.6 shows both diffusion and attention to the news at the global level, that is, when we aggregate all of our tweets and retweets across countries. In figure 1.6, the top-right panel

[9]These numbers are annualized according to the number of annual trading days and are net of the median risk premium in our full sample.

Table 1.3: Newspapers Dataset

Country	No. News Providers	Tweets	Retweets	Likes	Topics			
					Mortality	Quarant.	Med. Supply	Vaccines
Argentina	4	77,338	1,205,269	3,154,158	13%	10%	14%	63%
Australia	4	17,641	144,685	348,074	19%	39%	12%	29%
Brazil	4	32,552	1,331,641	8,705,983	45%	8%	15%	32%
Canada	5	48,545	442,675	861,102	33%	10%	17%	40%
Chile	4	33,954	408,095	630,825	56%	6%	10%	28%
China	3	32,791	947,706	2,580,177	39%	14%	19%	28%
Colombia	4	32,877	474,795	1,450,808	17%	12%	25%	45%
France	4	47,059	1,424,717	2,384,879	25%	26%	27%	22%
Germany	4	12,213	147,977	331,434	20%	24%	20%	35%
Hong Kong	3	21,204	419,964	606,520	17%	32%	21%	31%
India	4	103,907	937,468	5,611,874	32%	23%	16%	29%
Italy	3	33,696	265,658	715,064	10%	32%	29%	28%
Japan	4	18,984	156,845	277,461	18%	13%	30%	39%
Korea	4	13,459	82,445	143,649	44%	10%	26%	20%
Mexico	4	79,244	1,625,687	4,262,382	14%	11%	25%	50%
New Zealand	3	9,520	45,474	169,920	10%	40%	15%	35%
Spain	4	38,865	2,668,188	4,792,047	30%	20%	14%	36%
Switzerland	4	8,390	37,173	47,191	22%	20%	25%	33%
UK	4	25,356	1,145,404	2,286,491	27%	30%	15%	29%
USA	11	116,365	7,264,225	17,259,747	29%	7%	23%	41%
Total	84	803,960	21,176,091	56,619,786	26%	19%	20%	35%

Notes: This table shows summary statistics of COVID19-related news data that we collect for a large cross section of countries. Our real-time data range from 01/01/2020 to 02/22/2022. For each country, we report number of news providers and number of tweets collected. We also report the total number of retweets and likes as measures of attention. The last four columns report the share of tweets mentioning number of deaths, quarantine measures, medical supply, and vaccines, respectively.

of this figure provides a breakdown of the most prominent topics addressed in the COVID-19 tweets, namely, vaccines, death risk, quarantine measures, and availability of medical supplies. The attention to all of them increased substantially, with vaccines becoming prominent in the fall of 2020. In figure 1.6 bottom-panels, we document similar results for high-attention tweets, i.e., tweets ranked top-1% by the number of retweets within each one of our countries. For this subset of tweets, we also collected their retweets with a ‘quote’, i.e., with text written by the “retweeters" in order to study their tone. We find a high correlation between the tone of the original newspaper tweets and that of the quoted retweets, meaning that our methodology captures a relevant-and-consistent partition of tweets. We provide detailed results in the appendix (see table B.1).

Figure 1.7 shows the intraday pattern of the diffusion of COVID-19 news for each country. This figure is not based on universal time, rather it accounts for country-specific time. In each country, we consider two country-specific subsamples: the pre-epidemic period and the first wave of the epidemic. The first wave spans from the time a country records its 100th COVID-19 case to the peak of global daily deaths, on 01/18/2021. The pre-epidemic sample starts on 01/01/2020 and it ends on the time a country records its 100th COVID-19 case. There are two main takeaways from this picture: (i) the diffusion of COVID-19-related news increases

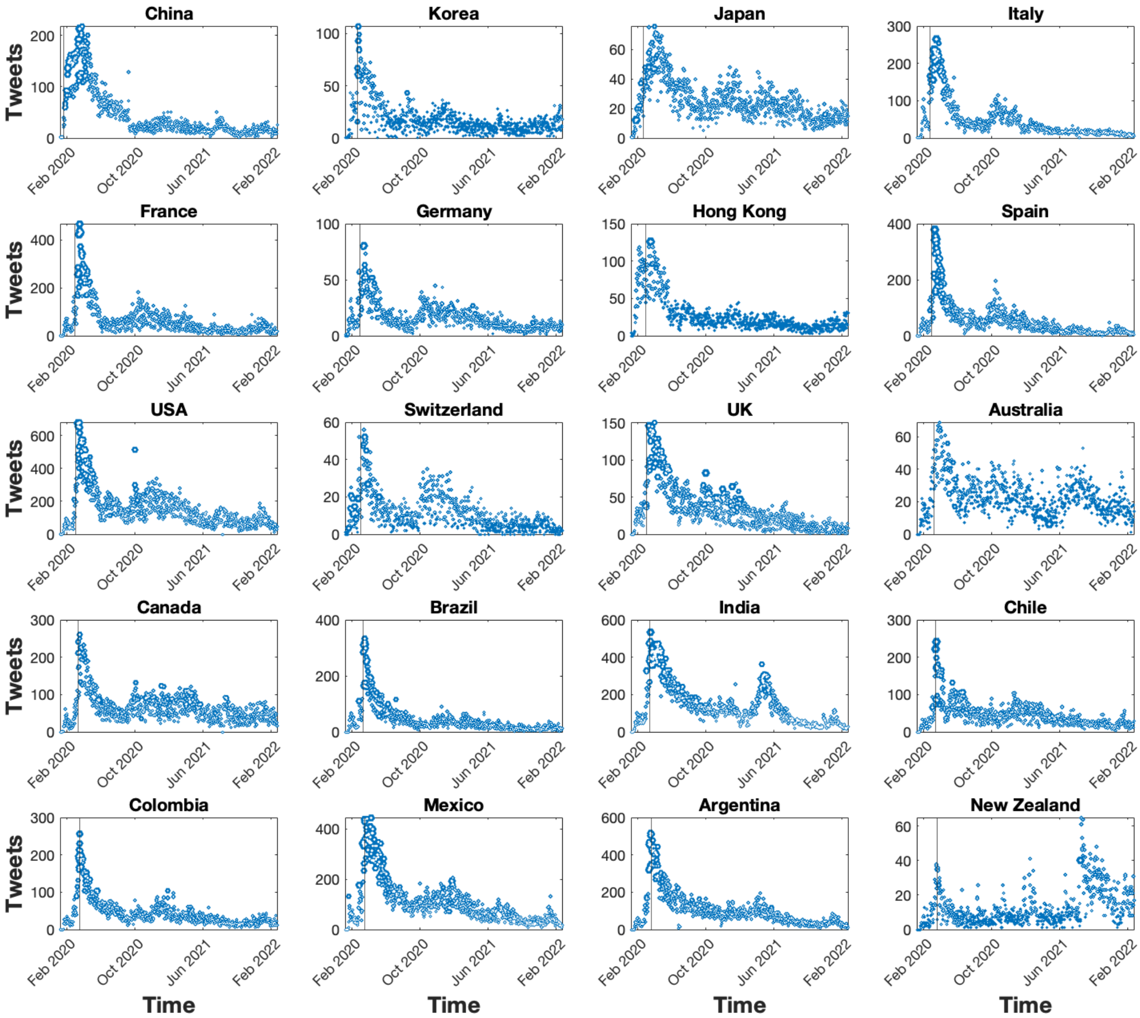


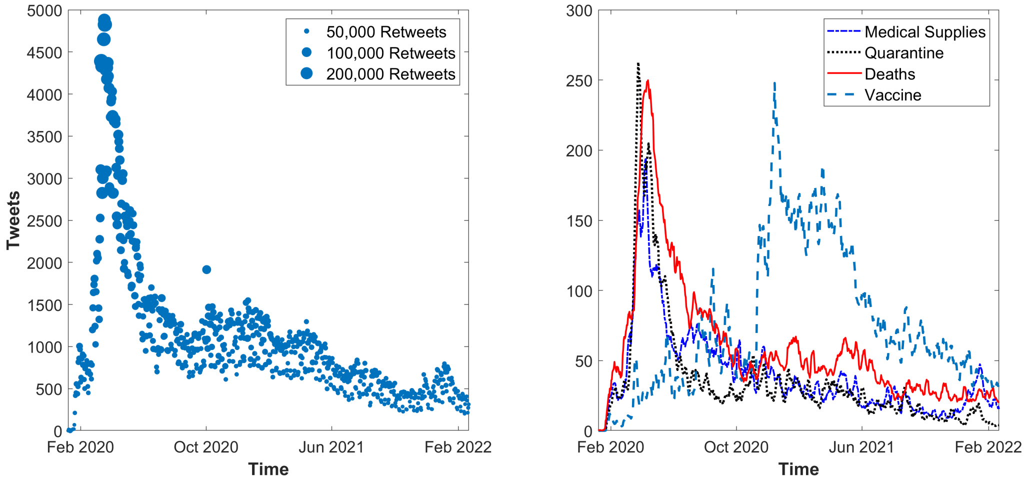
Figure 1.5: Information Diffusion and Attention across Countries

Notes: This figure shows the daily number of tweets posted in each country by major newspapers. The vertical axis shows the daily number of tweets. The size of each data point represents the number of retweets scaled by the maximum daily number of retweets for each country. The sample starts on 01/01/2020 and ends on 02/22/2022. The vertical line depicts the date that each country had more than 100 confirmed cases of COVID19. More details on the data collection are reported in the Appendix.

significantly with local epidemic conditions, and (ii) a significant share of the diffusion occurs while the local capital markets are open. Hence monitoring media activity can be a valuable tool for tracking the real-time information set of financial market participants.

Tweet Tone. Since we use Twitter activity to form a high-frequency risk factor, we need to identify the tone of the tweets, that is, we need to know whether they relate to either good or bad news. Given (i) the high volume of tweets that we collect and (ii) the fact that our tweets are written in different languages, we use Polyglot (available at https://pypi.org/project/polyglot/), i.e., a natural language pipeline that supports multilingual applications with polarity lexicons for 136 languages. This computer-based mapping algorithm

Panel A: Full Sample



Panel B: High-Attention Tweets

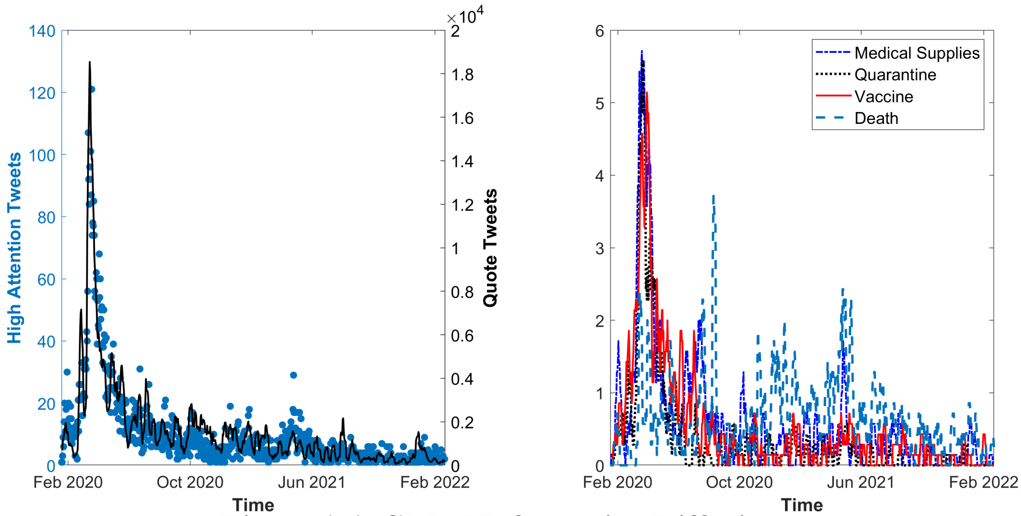


Figure 1.6: Global Information Diffusion

Notes: In panel a, the left panel of this figure shows the daily total number of tweets posted across countries by major newspapers. The vertical axis shows the daily number of tweets. The size of each data point represents the number of retweets scaled by the maximum daily number of retweets. The right panel shows the daily number of tweets related to death-risk, (scarcity of) medical supplies, quarantine, and vaccines. The tweets were identified using a multilingual bag-of-words approach. In panel b, we focus on high-attention tweets, i.e., top-1% by number of quote (re)tweets. The sample starts on 01/01/2020 and ends on 02/22/2022. More details on the data collection are reported in the Appendix.

reads our text and classifies the words into three degrees of polarity: +1 for positive words, -1 for negative words, and 0 for neutral words. We provide two examples in table A.2 (see our appendix).

Our measure of the tone of tweets is based on the count of positive words minus the count of negative words, divided by the sum of positive and negative word counts (Twedt and Rees, 2012). We compute this measure at the country level at both the hourly and the daily frequency,

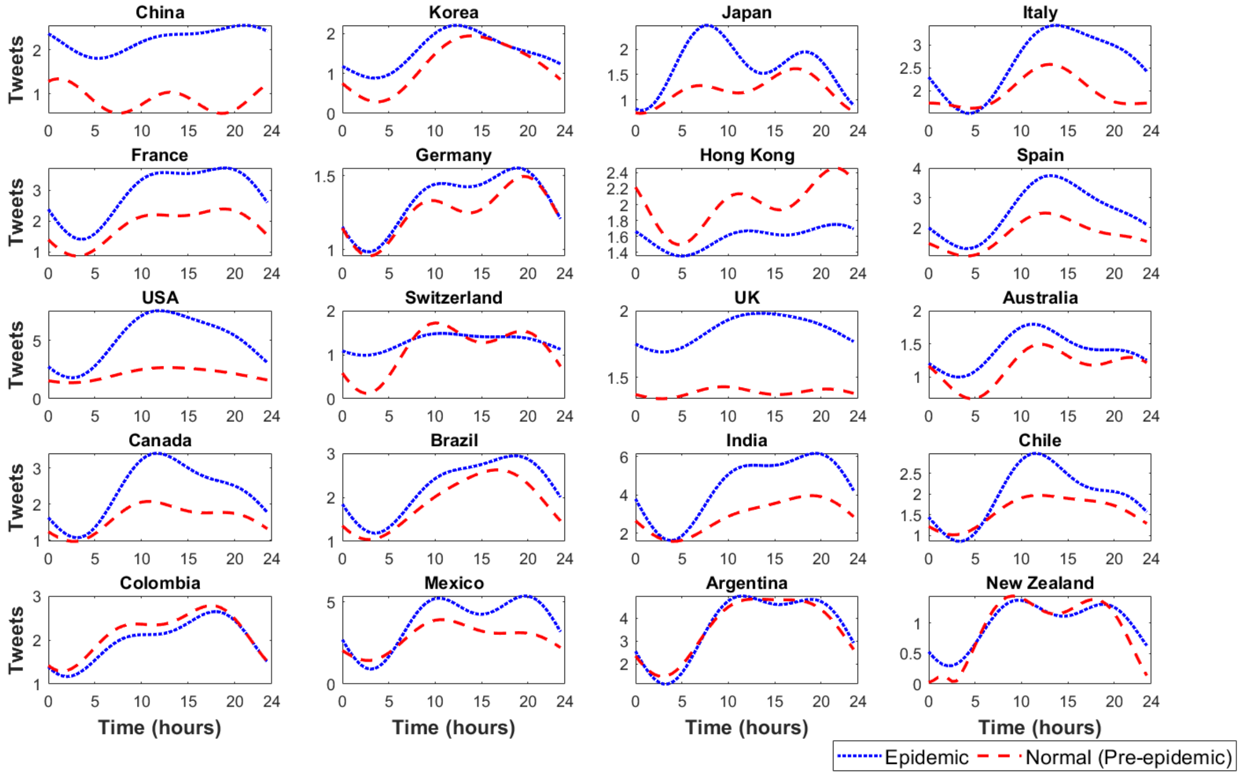


Figure 1.7: Intraday Information Diffusion

Notes: This figure shows the intra-day trend of the number of tweets posted every 30 minutes across several countries in our dataset. The dotted line depicts the trend during the first wave of the epidemic, while the dashed line shows the pre-epidemic trend. The first wave spans from the time a country records its 100th COVID-19 case to the peak of global daily deaths, on 01/18/2021. The pre-epidemic sample starts on 01/01/2020 and it ends on the time a country records its 100th COVID-19 case. The time reflects the local time zone of each newspaper. More details on the data collection are reported in the Appendix.

focusing solely on initial tweets and excluding any retweets. We then aggregate this measure across countries to obtain a global measure.

We depict our global tone factor in figure 1.8, left panel. Its time pattern is consistent with the observed contagion dynamics. Specifically, the tone became very negative by the end of January as the conditions in China started to precipitate. It improved in early February when there was still no sign of massive contagion in Europe, and it declined again when the epidemic started in Italy. The slow improvement of the tone of our tweets observed after the beginning of March pairs well with the observed flattening of the contagion curves in many of the countries in our dataset. We find these results reassuring as they confirm that our text analysis algorithm tracks the contagion dynamics in a reliable manner.

In addition, we note that collecting all original tweets and their retweets is computationally impossible for us. In table B.1 (see Appendix B, we show that there is a positive and significant correlation between the tone of the original tweets and that of the top-1% quote (re)tweets, meaning that our methodology captures a relevant-and-consistent partition of tweets.

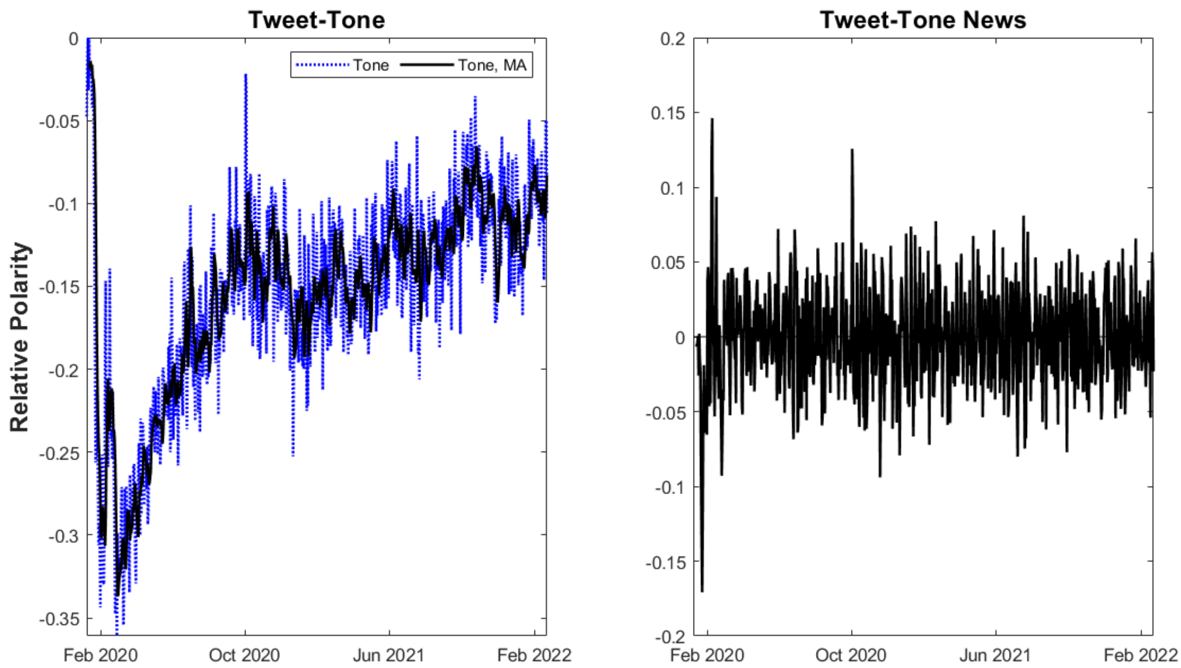


Figure 1.8: Twitter-Based COVID19 Factor

Notes: This figure shows our daily global Twitter-based COVID19 factor. We use Polygot to measure the polarity of our tweets and compute the tone of each tweet according to Twedt and Rees (2012). We aggregate the tones at a daily frequency and across countries. MA refers to a backward looking 5-day moving average. The news at time t is computed as the difference between the tweets-tone at time t and their MA at time $t-1$. The sample starts on 01/01/2020 and ends on 02/22/2022.

For the sake of our asset pricing analysis, we focus on the innovations to the tone of our tweets. One simple way to extract these innovations is to consider the difference in the tone at day t and its 5-day backward-looking moving average assessed at time $t-1$. We depict this time series in the right panel of figure 1.8 and note that it is nearly serially uncorrelated.

Cases and financial data. Cases data are from official medical bulletins. Our primary source is CSSE at Johns Hopkins University.[10] News to the contagion factor is obtained by computing the difference between the daily growth rate of contagion cases at time t and its backward-looking time $t-1$ moving average computed over the previous 5 days. We choose a 5-day window because it matches the number of days of a typical trading week.

Since our contagion-based factor spans a 7-day week, we assign to Friday the average growth rate of global contagion cases that occurred on Friday, Saturday, and Sunday.[11] Our financial data sources are detailed in table A.3 (see Appendix A). Our empirical asset pricing analysis takes into consideration the hypothesis that our countries may feature heterogeneous exposure

[10] https://github.com/CSSEGISandData/COVID19/tree/master/csse_covid_19_data/csse_covid_19_time_series

[11] For the Easter Holiday, we assign to Thursday the average daily growth rate of global cases from Thursday to the following Monday.

to global contagion risk.

Cases as a driver of time-varying exposure. One additional reason to focus on the number of cases as a relevant determinant of risk-premia is that many tweets in our sample focus on this topic. Specifically, we apply the Sievert and Shirley (2014) Latent Dirichlet Allocation (LDA) topic model to our covid-related tweets from English-written newspapers. Tweets are preprocessed, i.e., we remove stopwords and symbols such as #, we 'stem' our words, and account for both unigrams and bigrams. We apply the unsupervised machine learning model to our data at the country-level.[12] When λ is set to the canonical value of 0.5, in most of our English-speaking countries, the top unigrams and bigrams from the main topics include 'covid cases' or related terms. We report an example in figure B.1 (see Appendix B). Our data visualization webpage lets the interested reader choose different values of λ.

1.3.2 The Market Price of Viral Contagion News

Daily news. Every day, we group countries into three portfolios according to their relative number of COVID-19 cases measured the previous day. We do this separately for AEs and EEs. In Appendix B, we demonstrate that these results remain robust even when the share of COVID-19 cases is weighted by population (table B.2), indicating that our results are not driven by countries with larger populations. The H (L) portfolio comprises the top (bottom) countries in terms of COVID-19 cases. We also consider an investment strategy long in the H portfolio and short in the L portfolio. We refer to the returns of this portfolio as HML-COVID19.

We report common summary statistics for these portfolios in table 1.4. The turnover in each portfolio is moderate. The in-sample average of the returns in all portfolios is not different from zero, which is not surprising given our short sample, which comprises multiple contagion waves with natural peaks and valleys. All portfolio returns have substantial volatility and negative skewness. Focusing on the first quartile of the distribution of returns, we see that the portfolio comprising the more exposed countries tends to have more severe negative downside risk. This is an aspect that we capture in our conditional no-arbitrage model.

[12] Topics are indexed by $k = 1, ..., 5$. λ determines the weight given to the probability of term w under topic k relative to its lift (measuring both on the log scale). Setting $\lambda = 1$ results in the familiar ranking of terms in decreasing order of their topic-specific probability, and setting $\lambda = 0$ ranks terms solely by their lift.

Table 1.4: Summary Statistics for Portfolios

	Low	Medium	High	HML$_{COVID19}$
Panel A: Advanced economies				
Mean	0.014	0.039	0.020	0.006
	(0.060)	(0.060)	(0.072)	(0.031)
StDev	1.159	1.331	1.433	1.044
Skewness	-1.222	-0.762	-1.648	-0.108
First Quartile	-0.471	-0.496	-0.497	-0.545
Avg. N. Countries	5	4	5	-
Turnover (%)	0.5	1.3	0.6	-
Panel B: Emerging economies				
Mean	0.009	0.044	0.096**	0.086
	(0.087)	(0.094)	(0.049)	(0.063)
StDev	1.69	1.855	1.75	1.605
Skewness	-2.106	-1.255	-0.752	0.316
First Quartile	-0.662	-0.862	-0.774	-0.942
Avg. N. Countries	3	2	2	-
Turnover (%)	0.4	0.8	0.5	-

Notes: This table shows summary statistics for the equity excess returns of portfolios formed on a daily basis according to the relative share of country-specific COVID19 cases measured the day before formation. Hourly excess returns are in log units and multiplied by 100. Portfolios are obtained from equity indexes. Our real-time data range from 02/01/2020 to 02/22/2022. Turnover measures the number of countries entering or exiting a portfolio relative to the total number of countries in a specific portfolio $\times$ number of days in our sample. Numbers in parenthesis are HAC-adjusted standard errors.

Model and estimation. Given these preliminary observations, we consider the following conditional asset pricing model,

$$r^{ex}_{f,t+1} = \overline{r}^{ex}_{f,t} + \beta_{f,t} \cdot news^{glob}_{t+1}, \quad f \in \{H, M, L\}, \tag{1.8}$$

$$\beta_{f,t} = \beta_0 + \beta_{f,1} X_{f,t}, \tag{1.9}$$

$$\frac{\partial \overline{r}^{ex}_{f,t}}{\partial X_{f,t}} = \lambda \beta_{f,1}, \tag{1.10}$$

where $X_{f,t}$ is the share of contagion cases associated to portfolio f at time t, and λ is the market price of risk (MPR) of the global news factor $news^{glob}_{t+1}$.

This model can potentially capture many of the features of returns seen so far. First, it captures predictability through contagion-based time-varying betas. Second, it has the potential to capture higher negative skewness for countries that go through more severe contagion paths. Consider the case of portfolio H comprising countries receiving a sequence of relatively more severe adverse contagion news. This portfolio will have severe exposure to adverse news as the relative contagion share of the portfolio grows. When the relative contagion share starts to flatten out and decline, the sensitivity of this portfolio to good news is reduced ($|\beta_{H,t}|$ shrinks). This means that returns become less sensitive to positive news, and hence the right tail of the

Table 1.5: Summary of MPR estimation

	Covid Cases		Twitter News	
	A.E.	E.E.	A.E.	E.E.
Local units				
coef	-0.003^{***}	-0.006^{***}	0.013^{***}	0.007^{***}
se	(0.001)	(0.001)	(0.003)	(0.001)
USD units				
coef	-0.005^{***}	-0.005^{***}	0.011^{***}	0.006^{***}
se	(0.001)	(0.002)	(0.003)	(0.001)
Controlling for MKT				
coef	-0.002^{***}	-0.007^{***}	0.008^{***}	0.008^{***}
se	(0.001)	(0.001)	(0.002)	(0.001)

Notes: This table shows the results of the conditional linear factor model described in equations (1.8)–(1.10). Portfolios are formed on a daily basis according to the relative share of country-specific COVID19 cases measured the day before formation (X_t). On the left (right), the COVID19 factor is measured as the news to global COVID cases growth (tone of COVID-related tweets). When we measure the COVID19 news as unexpected number of contagion cases (unexpected improvement in COVID19-related tweets), we expect a negative (positive) market price of risk (MPR). Both daily excess returns and market prices of risk are in log units. Our cross section of test assets comprises both equity and bond portfolios. Our real-time data range from 02/01/2020 to 02/22/2022. Estimates and HAC-adjusted standard errors are obtained through GMM.

returns distribution is shortened.

Third, consistent with our previous descriptive returns, it accounts for heterogeneous exposure to global contagion news. Last but not least, it enables us to identify the market price of risk of this global contagion component, λ. By no-arbitrage, the extent of time-series predictability of our excess returns must equal $\lambda\beta_{f,1}$, and $\beta_{f,1}$ can be easily estimated in the time-series by considering the multiplicative factor $X_{f,t} \cdot news^{glob}_{t+1}$.

We report our main results obtained from daily data in table 1.5. The first two columns are based on unexpected changes in the growth of global contagion cases. The right-most columns are based on unexpected changes in the global tone of tweets. Note that the countries we consider provide daily updates about contagion cases at the end of the day. In order to properly represent the information set of investors, in our asset pricing model we lag the news by one day, i.e., we assume that day-t returns respond to news released in the evening of day $t-1$.

We estimate our asset pricing model through GMM and notice that all portfolios have an untabulated significant exposure to our contagion-based news, $\beta_{f,t}$.[13] In our sample, the portfolio of countries with the highest share of COVID-19 cases tends to be more exposed to contagion news. This sign is consistent with our expectations since positive (negative) news about global

[13]The share of contagion cases across our three portfolios have very different scales and variability. As a result, the coefficients $\beta_{f,1}$ are not revealing of the sorting of $\beta_{f,t}$ across portfolios. For this reason, we report only estimated MPRs.

contagion growth (tone of tweets) refers to an adverse shock to equity returns. Most importantly, the implied daily market price of risk is negative (positive) and significant with respect to contagion (tone of tweets) news. This means that the relative share of contagion cases forecasts an increase in expected future returns across all portfolios ($\lambda\beta_{f,1} > 0$). Equivalently, the share of contagion cases is a relevant positive predictor of the future cost of capital.

Our results hold regardless of whether we run our model using local currency returns or returns in USD. Furthermore, our results remain significant when we estimate a two-factor version of our model which controls for global market risk as measured by the MSCI Global Index.[14] Looking at the output of our specifications and accounting for estimation uncertainty, we conclude that 0.3% is a reasonable lower bound on the daily market price of risk of daily contagion news. We consider this estimate very significant, consistent with the great contraction experienced in equity markets during the first wave of the epidemic period.

Simultaneously, we note that this value is very plausible once we account for two observations. First, this is not the MPR of a financial factor and the associated estimated betas are very small. Second, contagion risk follows waves with a relatively short half-life. Equivalently, the exposure of our assets to this risk is small and relatively quick in reverting to zero.

Real-time estimation results. We visualize our results in figure 1.9. In contrast to what has been done so far, we estimate our model sequentially on samples of increasing length. Specifically, we start by estimating our model on a sample ending in June 2020, and then we re-estimate it by adding two weeks of data at the time. The very last estimation iteration delivers results identical to those reported in table 1.5. This exercise clarifies to which extent investors can identify exposures to contagion risk in real-time with our methodology.

In figure 1.9, panel (c), we show that our estimates of the MPR are statistically significant in almost the entire sample for both AEs and EEs. Since the figures are based on Covid cases, the MPR is negative. In figure 1.9 panels (a) and (b), we show the estimated risk premium on an HML-COVID19 strategy on either bond or equity portfolios across AEs and EEs. An HML-COVID19 strategy on bonds delivers a zero (negative) risk premium in AEs (EEs). Equities, instead, deliver a positive risk premium both in AEs and EEs. The estimated premium is smaller but better identified in AEs than in EEs. This result is important because it implies that containment policies that keep contagion cases relatively low may be very valuable in terms of saving lives and preventing severe financial wealth losses. The companion estimates of the exposure of these strategies can be found in figure B.2, Appendix B.

[14]Throughout our study when considering the MSCI index to control for the market, we use returns in USD. For EEs (EAs), we use the EE (EAs) MSCI.

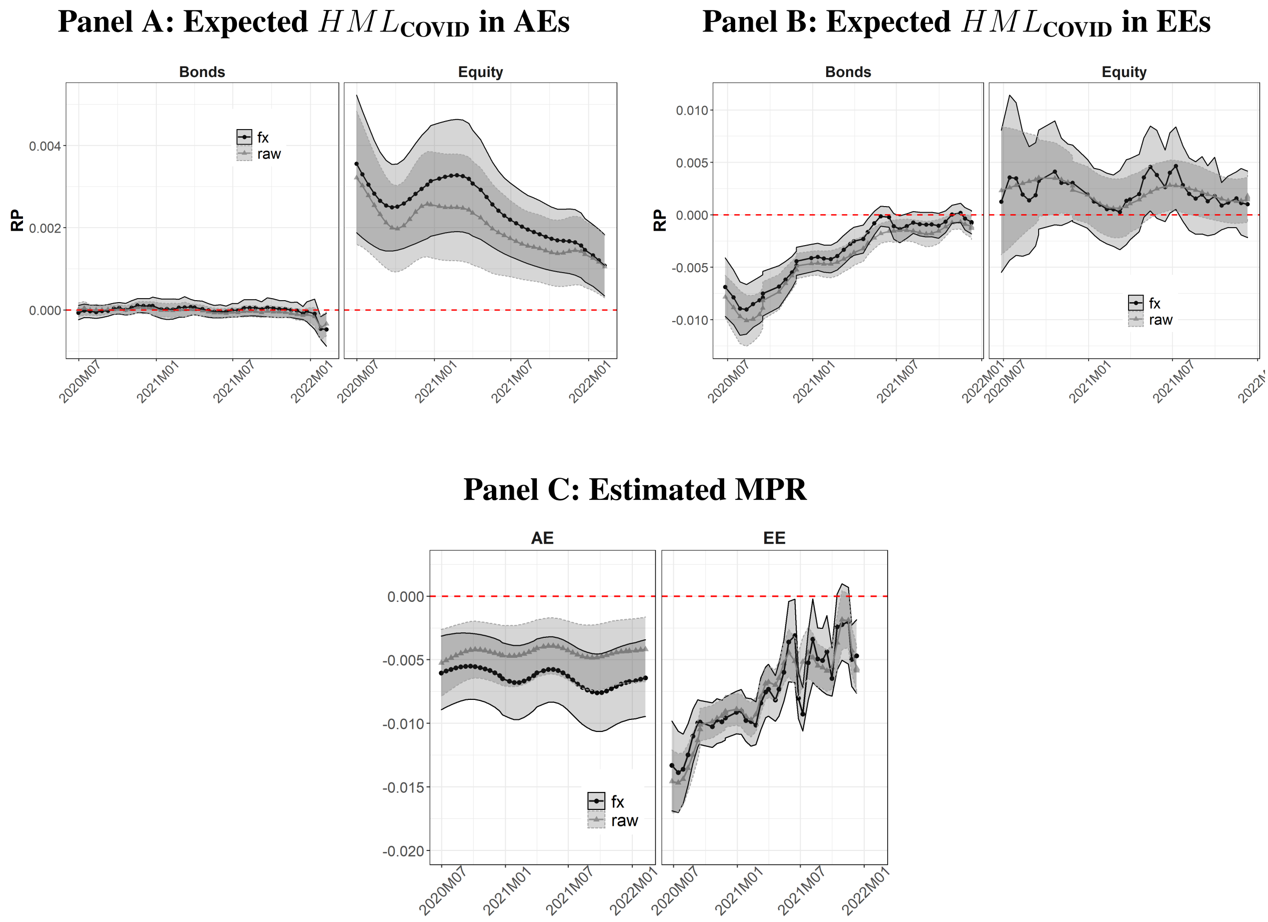


Figure 1.9: Sequentially Estimated MPRs and Risk Premia

Additional results with daily data. In table B.3 (see Appendix B), we show that replacing covid-related news with market returns in our conditional model delivers no positive and statistically significant market price of risk. This result confirms that (i) a conditional CAPM model fails in capturing viral contagion risk, and (ii) our measures are informative about viral risk.

K. French provides the FF5 factors at a daily frequency for developed countries (Fama and French (2017)). Given our limited cross section, estimating our model with time-varying betas for both our covid factor and the FF3/FF5 factors is not feasible. We take a hybrid approach and estimate a model in which the betas of our covid factor are time-varying, whereas the betas of the additional FF3/FF5 factors are constant. We report our estimated MPRs in table B.4 and confirm our main results.

So far, we have estimated a model with heterogeneous and time-varying exposure to a common risk factor related to global contagion news. Our dataset also enables us to construct AE- and EE-specific measures of COVID-19 case growth and Twitter tone. See, for example, figure B.3 in the Appendix.

We identify purely AE- and EE-specific components by regressing these fundamental measures on their global counterpart. The residuals of these two separate regressions represent AE- and EE-specific news for us. We show mixed results in Appendix B, table B.5 Specifically, when we use only equity-based test assets, local contagion news (panel A) is priced negatively in AEs and positively in EEs. Twitter-based local news (panel B) has a market price of risk statistically not different from zero. Local news is priced only when we use both bond and equity indices as test assets.[15] Given these considerations, we consider our specification with heterogeneous and time-varying exposure to global contagion risk news as more robust.

Controlling for jumps. Baker et al. (2020) point out that the COVID-19 epidemic period is characterized by a high frequency of extreme realizations of the US equity market returns. In order to check whether jump risk is driving our results, we repeat our analysis by using returns orthogonalized with respect to measures of jump risk. Specifically, we regress the equity returns of each one of the countries in our sample (indexed by i) on a country-specific daily dummy variable equal to $\texttt{sign}\, ret_{i,t}$ when $\texttt{abs}\, ret_{i,t} > 2.5\%$.

Since we work with a broad cross section of countries which includes also EEs, we consider an additional relative definition of jump risk. Namely, we construct a daily country-specific dummy variable equal to $\texttt{sign}\, ret_{i,t}$ when $\texttt{abs}\, ret_{i,t}$ is above the 99th percentile of its pre-epidemic country-specific distribution. We report our country-level percentiles in the appendix (table B.6 in the Appendix).

In table 1.6, we show our estimates of the market price of risk after controlling for realized jumps. Our main results continue to hold. In addition, we also consider a specification in which we orthogonalize our global factors with respect to a dummy variable that takes value $\texttt{sign}\, MSCI_t$ when the return of the MSCI global index is greater (smaller) than 2.5% (-2.5%). Our results confirm that pandemic risk carries a relevant market price.

Industry-level results. In order to further exploit our rich framework, we consider a more granular cross section by looking at industry-level equity indices in Europe. This task is relevant for at least two reasons: (i) industries are well known to be difficult to price; and (ii) our industries comprise firms based in different countries and hence their riskiness is based on an interesting mix of country-level contagion conditions.

Specifically, we utilize daily firm-level data for the entirety of the STOXX Europe 600, extracted from Bloomberg. The STOXX Europe 600 is a comprehensive index encompassing 600 constituents that span large, mid, and small capitalization companies across 17 European nations; we account for the quarterly compositional changes of the index. This selection represents nearly 90% of the free-float market capitalization in the European stock market. For our

[15] By “local news" we mean a variation in tone that is not spanned by a common global component. The content of the news may refer to either domestic or foreign events.

Table 1.6: MPR Estimation - Controlling for Jump Risk

	(1)		(2)		(1b)		(2b)	
	A.E.	E.E.	A.E.	E.E.	A.E.	E.E.	A.E.	E.E.
Local units								
coef	0.014^{***}	0.007^{***}	0.012^{***}	0.006^{***}	0.015^{***}	0.004^{***}	0.013^{***}	0.003^{***}
se	(0.002)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)
USD units								
coef	0.017^{**}	0.006^{***}	0.011^{***}	0.006^{***}	0.011^{**}	0.003^{***}	0.007	0.002^{**}
se	(0.007)	(0.001)	(0.003)	(0.001)	(0.005)	(0.001)	(0.005)	(0.001)
Controlling for MKT								
coef	0.012^{***}	0.007^{***}	0.006^{***}	0.007^{***}	0.012^{***}	0.006^{***}	0.007^{***}	0.004^{***}
se	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)

Notes: This table shows the estimates of the MPR for Twitter news controlling for jump risk. We pre-clean excess returns/factors by orthogonalize them with respect to an indicator variable equal to 1 when markets experienced a "jump". Except for this preliminary step, the portfolio formation, sample selection, and estimation procedure are the same as in table 1.5. In column (1) we orthogonolize the excess returns of our test assets with respect to a country-level daily dummy variable that takes the value $\texttt{sign}\, ret_{i,t}$ if $\texttt{abs}\, ret_{i,t} > 2.5\%$. In column (2), we repeat the same procedure but the dummy variable is equal to $\texttt{sign}\, ret_{i,t}$ if $\texttt{abs}\, ret_{i,t}$ is above the 99th percentile of its pre-epidemic country-specific distribution. In columns (1b) and (2b), we also orthogonalize our global risk factors using a global dummy variable that takes value $\texttt{sign}\, MSCI_t$ when the return of the MSCI global index is greater (smaller) than 2.5% (-2.5%).

analysis, we only focus on firms headquartered in one of the European countries in our sample (CH, DE, ES, FR, IT, SE, UK). We aggregate their equity returns at the industry level using the 2-digit Global Industry Classification Standard (GICS) code and obtain 11 distinct portfolios whose composition is detailed in table B.7. As shown in table 1.7, the market prices of risk estimated in our cross section of industries are comparable to those reported in table 1.5. This result supports our methodology as it shows that it can deliver significant results even at the industry-level, not just when pricing managed portfolio with country-level returns. In addition, in panel B and C we show the average exposure of each industry to our covid factors. Industries are ranked according to their average betas. We note two important points. First, the ranking of our industries when we define news as unexpected covid case growth is the perfect mirror image of the ranking when we use innovations to the tone of our Twitter news. Equivalently, once we account the the fact that the MPRs have opposite signs, the ranking is the same. We find this result reassuring.

Our estimation confirms that Energy, Utilities and Consumption Discretionary have been the most exposed to covid risk.[16] Health, Materials and IT, instead, have been relatively safer. This cross sectional average ranking is preserved in the time series through our sample.

[16]Consumption discretionary comprises segments that have been seriously armed during the pandemic such as Automobiles, Hotels, Restaurants and Leisure. A minority of segments, instead, have grown (for example, households durables and homefurnishing Retail)

Table 1.7: Industry-Level MPR Estimation

	Panel A: Market Price of Risk				
	Covid Cases		Twitter News		
	MPR Covid	MPR Kkt	MPR Covid	MPR Mkt	Obs
coef	-0.001^{**}	0.001^{***}	0.006^{***}	0.001^{**}	491
se	(0.000)	(0.000)	(0.000)	(0.000)	491

	Panel B: Average Covid Betas - Covid Cases			
Industry (i)	10 Energy	55 Utilities	25 Cons. Discr.	50 Commn.
$\text{Avg}_t\beta_{i,t}$	-2.092	-1.869	-1.684	-1.625
Industry (i)	40 Financials	30 Cons. Staples	60 Real Estate	20 Industrials
$\text{Avg}_t\beta_{i,t}$	-1.387	-1.375	-1.204	-1.189
Industry (i)	45 IT	15 Materials	35 Health	
$\text{Avg}_t\beta_{i,t}$	-1.103	-1.023	-0.926	

	Panel C: Average Twitter Betas - Twitter News			
Industry (i)	35 Health	15 Materials	45 IT	20 Industrials
$\text{Avg}_t\beta_{i,t}$	-0.126	-0.032	-0.002	0.034
Industry (i)	60 Real Estate	30 Cons. Staples	40 Financials	50 Commn.
$\text{Avg}_t\beta_{i,t}$	0.04	0.077	0.08	0.197
Industry (i)	25 Cons. Discr.	55 Utilities	10 Energy	
$\text{Avg}_t\beta_{i,t}$	0.251	0.34	0.443	

Notes: This table shows the results of the conditional linear factor model described in equations (1.8)–(1.10) applied to the industries detailed in table B.7. For each day t and industry i, we compute its exposure coefficient $\beta_{i,t}$ using the following industry-level share of contagion cases:

$$X_{i,t} = \sum_{f \in i} \omega_{f,t} X_{f,t},$$

where f denotes the firms in industry i, $X_{f,t}$ is the share of contagion cases in the country in which firm f is located, and $\omega_{f,t}$ is a value weight equal to the relative value of firm f with respect to the total value of the firms in industry i at time t. In panel A, the first (last) two columns refer to the results obtained when the COVID19 factor is measured as the news to global COVID cases growth (tone of COVID-related tweets). In all cases, we control for the market factor. Both daily excess returns and market prices of risk are in log units and expressed in EUR. In panel B (C), we report the average exposure coefficient for each industry when the factor is based on Covid cases (Twitter news). Our sample starts on 02/01/2020 and ends on 2/22/2022. Estimates and HAC-adjusted standard errors are obtained through GMM.

Intra-day news. An important advantage of our Twitter-based risk-factor is that we can measure it at very high frequencies, in contrast to daily contagion cases. Using higher frequency data may help sharpen the estimate of the market price of risk because it provides an increased number of observations.

This section focuses only on European countries whose markets are open simultaneously. Specifically, we focus on CH, DE, ES, FR, IT, SE and UK. Every day, we group them into three portfolios according to their relative number of COVID-19 cases. In table 1.8, we show our estimation results when we link hourly equity and bond excess returns to hourly Twitter-based news.

Table 1.8: Hourly Conditional Linear Factor Model

	β_0	$\beta_{L,1}$	$\beta_{M,1}$	$\beta_{H,1}$	MPR	N.Obs	N. Assets
			Panel A: equities and bonds, equities betas				
Hourly log returns							
coef	-0.090^{***}	9.879^{***}	4.043^{***}	2.853^{***}	0.014^{***}	4190	6
se	(0.007)	(0.712)	(0.294)	(0.207)	(0.003)	4190	6
Hourly log EUR returns (adjusting for FX)							
coef	-0.083^{***}	9.164^{***}	3.773^{***}	2.673^{***}	0.017^{***}	4190	6
se	(0.006)	(0.598)	(0.249)	(0.177)	(0.003)	4190	6
Hourly log returns controlling for the Market							
coef	-0.158^{***}	16.892^{***}	6.980^{***}	4.968^{***}	0.009^{***}	3951	6
se	(0.014)	(1.549)	(0.643)	(0.457)	(0.003)	3951	6
			Panel B: equities and bonds, bond betas				
Hourly log returns							
coef	-0.062^{***}	6.872^{***}	2.780^{***}	1.966^{***}	0.014^{***}	4190	6
se	(0.005)	(0.496)	(0.201)	(0.144)	(0.003)	4190	6
Hourly log EUR returns (adjusting for FX)							
coef	-0.058^{***}	6.385^{***}	2.609^{***}	1.851^{***}	0.017^{***}	4190	6
se	(0.004)	(0.421)	(0.174)	(0.124)	(0.003)	4190	6
Hourly log returns controlling for the Market							
coef	-0.109^{***}	11.743^{***}	4.831^{***}	3.439^{***}	0.009^{***}	3951	6
se	(0.010)	(1.072)	(0.442)	(0.315)	(0.003)	3951	6

Notes: This table shows the results of the conditional linear factor model described in equations (1.8)–(1.10). Portfolios are formed on a daily basis according to the relative share of country-specific COVID19 cases measured the day before formation (X_t). The coefficient $\beta_{f,t} = \beta_0 + \beta_f X_{f,t}$ refers to the exposure of the equity portfolio $f \in \{H, M, L\}$ to the COVID19 factor. We measure hourly COVID19 news as unexpected improvement in the hourly tone of COVID19-related tweets. Both hourly excess returns and market prices of risk are in log units. When we control for the market, returns are in USD, the market is measured by the MSCI Global Index and our factor model comprises a total of two factors. Our real-time data range from 02/01/2020 to 02/22/2022. Estimates and HAC-adjusted standard errors are obtained through GMM.

As for daily data, we consider multiple specifications of our no-arbitrage model. In this case, we also report our estimated beta coefficients. The implied market price of risk is positive, well-identified, and sizable. Our implied betas continue to be positive, i.e., viral contagion is priced as a source of risk. Consistent with the failure of the international-CAPM documented in table B.3, our the implied market price of risk is still positive and sizable when we control for the market and use a broader cross section of test assets.

There may be several country-specific characteristics (e.g., fiscal conditions, competition, etc.) that could make our portfolios differently exposed to pandemic risk–in order to address this concern, we replace β_0 in equation (1.9) with β_0^f, i.e., a country-specific fixed effect to its exposure. Thanks to the hourly frequency, we have enough observations to estimate this richer model. Our results continue to hold and are reported in table B.8.

Table 1.9: Vol-Adjusted Conditional Linear Factor Model

	β_0	$\beta_{L,1}$	$\beta_{M,1}$	$\beta_{H,1}$	MPR	N.Obs	N. Assets
			Panel A: equities, news from Twitter				
Daily log returns							
coef	-0.312^{***}	41.132^{***}	19.797^{***}	6.102^{***}	0.006^{**}	499	3
se	(0.053)	(5.775)	(2.585)	(1.468)	(0.003)	499	3
			Panel B: equities, news from Twitter				
Hourly log returns							
coef	-0.005^{***}	0.453^{***}	0.174^{***}	0.134^{***}	0.054^{**}	4301	6
se	(0.001)	(0.131)	(0.055)	(0.040)	(0.028)	4301	6

Notes: This table shows the results of the conditional linear factor model described in equations (1.8)–(1.10). Portfolios are formed on a daily basis according to the relative share of country-specific COVID19 cases measured the day before formation (X_t). The coefficient $\beta_{f,t} = \beta_0 + \beta_f X_{f,t}$ refers to the exposure of the equity portfolio $f \in \{H, M, L\}$ to the COVID19 factor. We measure hourly (daily) COVID19 news as unexpected improvement in the hourly (daily) tone of COVID19-related tweets. We project this factor on realized market volatility and use the implied residual in our estimation. Both excess returns and market prices of risk are in log units and are expressed in USD. The market is measured by the MSCI Global Index. Our real-time data range from 02/01/2020 to 02/22/2022. Estimates and HAC-adjusted standard errors are obtained through GMM.

Controlling for Volatility. In this last step of our research, we project our Twitter-based COVID factor on realized market volatility and use the implied residual to redo our analysis. Equivalently, we look at COVID news that are orthogonal to pure volatility shocks. We measure realized volatility as the standard deviation of the MSCI Global Index at the daily (hourly) frequency using minute-level data. We report our results in table 1.9. Both daily data and intraday data confirm that contagion news has an extremely high MPR, even after controlling for volatility.

International Flows. To further validate our results, we study international investment flows related to the countries in our cross section. Weekly net flows are from EPFR and they are rescaled by country-level GDP so that our results are not driven by country size. In this step, we exclude the US, given its unique role in international markets (among others, see Maggiori (2017)). Over such a long span of time, capital flows toward/from the US were driven by many other factors above and beyond COVID. After forming portfolios according to relative contagion levels, we forecast one-week ahead flows using the (lagged) weekly share of portfolio-level COVID-19 cases.

As reported in table 1.10, countries that start the week with a higher level of relative contagion are expected to receive lower net inflows ($\beta_1 < 0$)). This effect is reversed ($\beta_1 > 0$)) when we focus on net bond flows in EE, consistent with the idea that they may be perceived as safer assets and hence their demand may actually increase due to flight to safety.[17] We visualize

[17]Recall that in EEs, bonds provide insurance against bad covid news (see figure 1.9).

Table 1.10: International Flows and News

	Bonds		Equities	
	AE	EE	AE	EE
β_0	0.247^{***}	-0.171^{***}	0.589^{***}	0.135^{***}
	(0.026)	(0.032)	(0.051)	(0.036)
β_1	-0.921^{***}	0.311^{***}	-4.500^{***}	-0.996^{***}
	(0.073)	(0.070)	(0.274)	(0.123)
J-stat	11.234	11.825	7.069	11.676
N	75	71	75	71

Notes: This table reports the results of the following linear system:

$$FL_t^f = \beta_0 + \beta_1 X_{t-1}^f + \epsilon_t^f$$

where FL_t^f is the flow to funds that invest in portfolio $f \in \{H, M, L\}$ during week t rescaled by portfolio$-f$ 2019 GDP; X_{t-1}^f refers to the weekly share of portfolio-specific COVID19 cases. Portfolios are formed on a weekly basis according to the relative share of country-specific COVID19 cases measured the week before formation. Fund flows-to-GDP is expressed in basis points (bps). Our data range from 02/01/2020 to 07/14/2021 at a weekly frequency. Estimates and HAC-adjusted standard errors are obtained through GMM.

these findings in figure B.4, Appendix B.

1.4 Conclusion

In this study, we quantify the exposure of major financial markets to news shocks about global contagion risk, taking into account local epidemic conditions. We construct a novel data set that includes (i) medical announcements related to COVID-19 for a broad cross section of countries; and (ii) high-frequency data on epidemic news diffused through Twitter. Across several classes of financial assets and currencies, we provide novel empirical evidence about financial dynamics surrounding epidemic announcements, both at a daily frequency and an intra-daily frequency. Formal estimations based on both contagion data and social media activity about COVID-19 confirm the significant market price of epidemic risk. We conclude that policies related to the prevention and containment of contagion could be precious not only in terms of lives saved but also in terms of preserving global financial wealth. Future research should study the interplay of our analysis and the methodology in Diercks et al. (2023).

Bibliography

Acharya, V. V., and S. Steffen. 2020. The risk of being a fallen angel and the corporate dash for cash in the midst of covid. *CEPR's Covid Economics Review No. 10*.

Ai, H., and R. Bansal. 2018. Risk preferences and the macroeconomic announcement premium. *Econometrica* 86(4).

Ai, H., R. Bansal, and L. J. Han. 2022a. Information acquisition and the pre-announcement drift. *Working Paper*.

Ai, H., L. J. Han, X. N. Pan, and L. Xu. 2022b. The cross section of the monetary policy announcement premium. *Journal of Financial Economics* 143(1).

Albuquerque, R. A., Y. J. Koskinen, S. Yang, and C. Zhang. 2020. Love in the time of covid-19: The resiliency of environmental and social stocks. *CEPR's Covid Economics Review No. 11*.

Augustin, P., V. Sokolovski, M. G. Subrahmanyam, and D. Tomio. 2021. In sickness and in debt: The covid-19 impact on sovereign credit risk. *Journal of Financial Economics*.

Baker, S. R., N. Bloom, and S. J. Davis. 2016. Measuring economic policy uncertainty*. *The Quarterly Journal of Economics* 131(4):1593.

Baker, S. R., N. Bloom, S. J. Davis, K. Kost, M. Sammon, and T. Viratyosin. 2020. The unprecedented stock market reaction to covid-19. *The Review of Asset Pricing Studies* 10(4).

Bianchi, F., R. Cram, and H. Kung. 2021. Using social media to identify the effects of congressional viewpoints on asset prices. *Working Paper*.

Bonaccolto, G., N. Borri, and A. Consiglio. 2019. Breakup and default risks in the eurozone. *Available at SSRN 3487453*.

Bretscher, L., A. Hsu, P. Simasek, and A. Tamoni. 2020a. Covid-19 and the cross-section of equity returns: Impact and transmission. *The Review of Asset Pricing Studies* 10(4).

Bretscher, L., A. Hsu, and A. Tamoni. 2020b. The supply channel of uncertainty shocks and the cross-section of returns: Evidence from the covid-19 crisis. *Available at SSRN 3588418*.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu. 2020a. The structure of economic news. *National Bureau of Economic Research*.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu. 2020b. Business News and Business Cycles. *Working paper*.

Carletti, E., T. Oliviero, M. Pagano, L. Pelizzon, and M. G. Subrahmanyam. 2020. The covid-19 shock and equity shortfall: Firm-level evidence from italy.

Cororaton, A., and S. Rosen. 2020. Public firm borrowers of the us paycheck protection program. *CEPR's Covid Economics Review No. 15*.

Da, Z., J. Engelberg, and P. Gao. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2011.01679.x 2011. In search of attention. *The Journal of Finance* 66(5):1461–1499. doi: 10.1111/j.1540-6261.2011.01679.x.

Darmouni, O., K. Siani, and K. Xiao. 2023. Nonbank fragility in credit markets: Evidence from a two-layer asset demand system. *NBER Working Paper*.

Diercks, A. M., A. Hsu, and A. Tamoni. 2023. When it rains it pours: Cascading uncertainty shocks. *Journal of Political Economy*.

Engle, R. F., S. Giglio, B. Kelly, H. Lee, and J. Stroebel. 2020. Hedging Climate Change News. *The Review of Financial Studies* 33(3).

Epstein, L., and S. E. Zin. 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57:937–969.

Fama, E. F., and K. R. French. 2017. International tests of a five-factor asset pricing model. *Journal of Financial Economics* 123(3).

Garmaise, M., Y. Levi, and H. Lustig. 2021. Spending less after (seemingly) bad news. *Working paper*.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus. 2020. Inside the mind of a stock market crash. *arXiv preprint arXiv:2004.01831*.

Goldstein, I., R. Koijen, and H. Mueller. 2021. Covid-19 and its impact on financial markets and the real economy. *The Review of Financial Studies*.

Gormsen, N. J., and R. S. J. Koijen. 2020. Coronavirus: Impact on stock prices and growth expectations. *The Review of Asset Pricing Studies* 10(4):574–597.

Gormsen, N. J., R. Koijen, and I. Martin. 2021. Implied dividend volatility and expected growth. *Working Paper*.

Han, L. J. 2020. Iannouncements, expectations, and stock returns with asymmetric information. *Working Paper*.

Hartley, J. S., and A. Rebucci. 2020. An event study of covid-19 central bank quantitative easing in advanced and emerging economies.

Hassan, T. A., S. Hollander, L. Van Lent, and A. Tahoun. 2019. Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics* 134(4):2135–2202.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun. 2020. Firm-level exposure to epidemic diseases: Covid-19, sars, and h1n1. *Working Paper Series*.

Kaniel, R., and P. Wang. 2020. Unmasking mutual fund derivative use during the covid-19 crisis. *CEPR's Covid Economics Review No. 49*.

Koijen, R. S. J., and M. Yogo. 2019. A demand system approach to asset pricing. *Journal of Political Economy* 127(4).

Lucca, D. O., and E. Moench. 2015. The pre-fomc announcement drift. *The Journal of Finance* 70(1).

Maggiori, M. 2017. Financial Intermediation, International Risk Sharing, and Reserve Currencies. *American Economic Review* 107(10).

Manela, A., and A. Moreira. 2017. News implied volatility and disaster concerns. *Journal of Financial Economics* 123(1):137–162.

Papanikolaou, D., and L. D. W. Schmidt. 2020. Working remotely and the supply-side impact of covid-19. *Working Paper*.

Pástor, L., and B. Vorsatz. 2020. Mutual fund performance and flows during the covid-19 crisis. *The Review of Asset Pricing Studies* 10(4).

Ramelli, S., and A. Wagner. 2020. Feverish stock price reactions to covid-19. *Forthcoming, Review of Corporate Finance Studies; Swiss Finance Institute Research Paper No. 20-12*.

Schmeling, M., and C. Wagner. 2023. oes central bank tone move asset prices? *Journal of Financial and Quantitative Analysis*.

Sievert, C., and K. Shirley. June 2014. LDAvis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*.

Sinagl, P. 2020. Cash-flow risk and industry response to covid-19 outbreak. *Available at SSRN 3566511*.

Twedt, B., and L. Rees. 2012. Reading between the lines: An empirical examination of qualitative attributes of financial analysts' reports. *Journal of Accounting and Public Policy* 31(1):1–21.

van Binsbergen, Jules H. Bryzgalova, S., M. Mukhopadhyay, and V. Sharma. 2022. (almost) 200 years of news-based economic sentiment. *Working Paper*.

Chapter 2

International Climate News

Co-authored with R. Colacito, M.M. Croce, and B. Yang.[1]

Abstract

We develop novel high-frequency indices that measure climate attention across a wide range of developed and emerging economies. By analyzing the text of over 23 million Tweets published by leading national newspapers, we find that a country experiencing more severe climate news shocks tends to see both an inflow of capital and an appreciation of its currency. In addition, brown stocks experience large and persistent negative returns after a global climate news shock if located in highly exposed countries. A risk-sharing model in which investors price climate news shocks and trade consumption and investment goods in global markets rationalizes these findings.

Keywords: International Climate Indices, Text Analysis, Trade, Currencies.
JEL Codes: F3, F4, G1.

[1]Full text and Internet Appendix available at SSRN.

2.1 Introduction

In recent years, there has been a growing focus on the interaction between climate dynamics and economic activities. Predicted changes in climate patterns are expected to result in a significant redistribution of economic operations across various locations, sectors, and corporations. At the same time, the valuation of stocks and bonds is increasingly reflecting climate risk factors. However, to the best of our knowledge, there has been limited attention given to the relationship between climate-related news shocks and international economic variables, such as currencies and global capital flows.

In this paper we construct a new index that captures high-frequency climate attention based on the content of the Twitter feeds of major newspapers in both advanced and emerging economies. We document two main findings. First, countries with greater exposure to climate news shocks experience a persistent currency appreciation and a decrease in net exports, highlighting the significant impact of climate news on economic variables. Second, brown firms located in highly exposed countries tend to experience large negative and persistent stock returns after adverse climate news shocks. To the best of our knowledge, we are the first to document no significant effect on high-emission firms located in countries with low exposure to climate news shocks.

Specifically, we employ textual analysis techniques to analyze news articles from major national newspapers posted on Twitter during the period of 2014–2022. In order to form hypotheses on the response of currencies, we start with a flexible no-arbitrage model with complete markets in which (i) climate news shocks are priced, and (ii) countries feature heterogeneous exposure to global climate news shocks. This setting suggests that countries with high exposure to climate news shocks should experience an appreciation of their currency upon the arrival of adverse global news shocks. When testing this hypothesis using the high-frequency dataset that we constructed, we found evidence that corroborates our prediction. This result adds empirical support to the theoretical link between climate news shocks and currency values.

Additionally, in line with the empirical asset pricing literature, we estimate sensitivity coefficients of returns over various horizons and for more than 17,000 firms. We examine firms with varying emission intensities located in all the countries included in our study. Importantly, in our empirical investigation we use an interaction term comprising emissions and country-level exposure. The literature has already documented that climate attention shocks are followed by negative cumulative stock returns. We find that this effect is highly persistent and is concentrated among brown firms that are located in countries with high exposure. In contrast, brown firms in countries with low exposure exhibit no significant response in their equity returns.

In a second step, we consider a general equilibrium model in which: (i) investors price climate news shocks; (ii) investment goods can be used to increase either green or brown assets; and

(iii) there are both productivity and global climate news shocks. This model predicts that a country subject to a relatively more adverse climate news shock should experience both an appreciation of its currency (as in the no-arbitrage model) and a decline in its net exports. Our newly developed indicators can be aggregated at a quarterly level, enabling us to merge our news-based data with international trade data. Our results provide empirical validation for our hypotheses.

In what follows, we provide a more detailed description of the novel contribution related to our climate news index. More specifically, as in Engle et al. (2020), we construct a climate attention index that measures the extent to which climate change is discussed in the news media. However, our method differs in that we focus on newspapers with a significant presence on Twitter across many countries. This approach allows us to compile a large dataset of over 23 million tweets, which we then aggregate at the country level across daily, weekly, monthly, and quarterly frequencies. We compare the aggregated text to a corpus of authoritative texts on climate change, similar to the method used by Engle et al. (2020). As of the date of this draft, our data coverage refers to a total of 25 countries that span a wide range of local languages, income levels, and geographical regions.

We construct climate news indices at the country level and aggregate them to create a global index by computing cross-country averages weighted equally, by GDP, and by Twitter volume (data available on our website). For each country in our dataset, we regress its country-specific climate index on the global index. The estimated slope, denoted as β, measures each country's exposure to common or, equivalently, global news. Across countries, we find a substantial degree of variation in exposure.

To further assess our text-based indices, we assess the correlation between our estimated β values and alternative measures of climate exposure. We first use the vulnerability score provided by the Notre Dame Global Adaptation Initiative. This comprehensive index aims to encapsulate "a country's exposure, sensitivity, and capacity to adapt to the negative effects of climate change." We observe a positive and statistically significant correlation between our β estimates and this index, implying that countries with higher β values exhibit greater susceptibility to climate change.

In addition, we also consider absolute latitude as a second indicator of climate exposure. Countries situated in regions with lower absolute latitudes are usually considered more exposed to rising temperatures resulting from climate change. We find a negative correlation between our β estimates and absolute latitude, reinforcing the effectiveness of our approach in capturing a country's exposure to climate change. Importantly, we show that our measure is correlated with vulnerability but not fully subsumed by it, indicating that our exposure coefficients provide additional information beyond what is captured by vulnerability.

Leveraging our granular dataset, we use high-frequency data to identify the impact of climate news shocks on asset prices. Specifically, we analyze the reaction of currencies and stocks in the days following a substantial increase in climate attention. To achieve this, we compute daily innovations in our global climate index and focus on the top 15% of these observations. This approach allows us to (i) reduce noise from non-climate-related news and (ii) better isolate days driven by "pure" climate news shocks. Such a methodology is commonly employed in the literature on identifying the effects of news related to policy announcements (Leombroni et al. 2021). Importantly, some days with substantial climate-related news might fall below the top 15% threshold due to the prominence of other topics in the news. Consequently, as is typical in this literature, we may have a downward bias, i.e., a reduced likelihood of detecting a significant effect of climate news on asset prices and currencies.

We further incorporate a sentiment analysis of climate-related tweets on days with high climate attention. Using the BERTopic algorithm, we cluster all tweets into topics and identify those closely related to climate. This subset is then divided into two groups based on whether the topic primarily pertains to climate-related physical or transition risk. Following this classification, each tweet on a given day is categorized as related to physical risk, transition risk, or non-climate topics. We apply multilingual sentiment classifiers to distinguish between positive and negative content. Our findings indicate that days with extremely high climate attention are predominantly characterized by negative climate-related news. The reaction of currencies and stock returns is mainly driven by negative news about transition risk.

Our analysis using daily-frequency data reveals that currencies of countries with higher exposure to our global climate index tend to appreciate following climate news shocks, with this effect persisting over several days. Remarkably, this finding is observed not only in pairs of advanced economies but also in combinations involving at least one developing country. Moreover, the robustness of this finding is upheld regardless of (i) the method used for aggregating country-level climate indices into a single global index, and (ii) whether we use weekly or daily data.

In the second part of the paper, we develop a general equilibrium model that incorporates a recursive risk-sharing mechanism for both productivity and climate news shocks. To the best of our knowledge, this is the first general equilibrium model in the climate finance literature to incorporate priced climate news and multiple countries. While the model is necessarily stylized to focus on the key mechanism, it provides a novel and valuable framework for understanding the role of this type of risks in an international context. The significance of climate news shocks lies in their impact on the distribution of climate-related utility damages among countries, which are characterized by heterogeneous degrees of exposure to a shared climate news shock. Within this framework, our model demonstrates that when agents price negatively climate news shocks, two outcomes occur in response to an adverse global news shock: the currency of the most exposed country appreciates, and there is a decrease in its net exports, signifying a resource outflow from the least exposed country. We proceed to empirically test these theoretical outcomes using our data set.

We test the predictions of our model by using quarterly trade data combined with our climate index aggregated at the same frequency. As in the model, we find that countries more exposed to climate news shocks tend to become net recipients of capital following adverse climate news, as evidenced by a decrease in their net exports.

Related literature. An important strand of recent literature employs textual analysis of earnings conference call transcripts to assess exposure to climate risk (see, for example, Sautner et al. 2023, Hassan et al. 2019 and Hassan et al. 2023). We differ from this literature in that we focus on the public interest for climate, rather than on educated firm-level experts. Moreover, our access to social media outlets allows us to construct a high-frequency index for a large cross-section of countries.

Our frictionless model builds on the tradition of international macro-finance models with recursive preferences and news shocks (Colacito et al. 2018). We present a unified framework to document that currencies, international capital flows, and investments in brown and green technologies are related to our climate-related news. We leave the study of settings with segmented markets (see, among others, Sandulescu et al. 2020) to future work.

A substantial body of literature has examined the behavior, predictability, and volatility of exchange rates (Della Corte et al. 2009; Della Corte et al. 2011; Della Corte et al. 2016). Our analysis contributes to this stream by focusing on the dynamic response of the foreign exchange market to news. In international asset pricing, extensive research has developed and analyzed models to understand asset prices, exchange rates, and macroeconomic interactions across countries (Pavlova and Rigobon 2007, 2010, 2013; Hassan 2013; Hassan et al. 2015, 2016; Zviadadze 2017). We contribute to the broader understanding of international macroeconomic dynamics and asset valuation in response to global climate-related events.

Another line of research investigates the impact of political and economic risk on firms, asset markets, and international economic linkages (Hassan et al. 2019, 2023). This literature develops methodologies to quantify various risks and uncertainties and assesses their economic and financial spillovers. Our study introduces a novel text-based measure of global attention to climate change as an indicator of perceived risk and uncertainty. Furthermore, we document the influence of our index on currency markets and international financial dynamics.

The literature on carry trade strategies has linked their risk premia to global risk factors (see, for example, Lustig et al. 2011, 2014; Mueller et al. 2017). Our results suggest that climate-related global news is an important risk factor for both exchange rate movements and capital flows.

Lee et al. (2022) find heterogeneous impulse responses of monthly U.S. dollar (USD) real exchange rates of 76 countries to global temperature shocks. We differ from their work in several dimensions: (i) we have a broader cross-section of country pairs; (ii) our climate-related indices move with a broad set of climate-related news, not just realized temperature changes; and (iii) we link both currencies and international current accounts. Javadi et al. (2023) find no link between G10 currencies and the vulnerability of these countries to physical climate risk. Our broader indicator captures both physical and transition risk and suggests the presence of a relevant link at both the daily and the weekly frequency.

Our work is related to the evolving academic field of climate finance (see, among others, Giglio et al. 2021a,b, 2023; Starks 2023; Hsu et al. 2023; Bansal et al. 2019; Pástor et al. 2021, 2022; Colacito et al. 2019; Acharya et al. 2022; Bolton and Kacperczyk 2023, 2021; Bolton et al. 2023; and Kashyap et al. 2024). The climate macro-finance literature has primarily focused on the link between climate risks, local assets and local macroeconomic dynamics. We differ from these studies for our attention to international quantities and currencies.

2.2 Data

Data collection. In line with the methodology developed in Engle et al. (2020), we construct a climate attention index to assess the extent of discourse concerning climate change within the news media. Our approach differs from that of Engle et al. (2020) in that we have chosen to focus on news media outlets, particularly newspapers, that have a strong presence on the social media platform Twitter.[2] Utilizing textual analysis, we generate country-level indices at both low and high frequencies. As of the date of this draft, our dataset encompasses a total of 25 countries, providing a comprehensive representation across a wide spectrum of local languages, income levels, and geographical regions. We have conveniently summarized our dataset in Table 2.1.

Climate attention index construction. We use the "universe" of Tweets from October 2014 until December 2022 from the Twitter accounts of official newspapers (the full list of newspapers is reported in table A.1 in the appendix). We construct our index by comparing the news content in these tweets to a corpus of "[...] authoritative texts on the subject of climate change" as in Engle et al. (2020). Similar to Engle et al. (2020), we aggregate the authoritative text documents into a Climate Change Vocabulary (CCV), which amounts to the list of both unique terms and their associated frequency with which each one of them appears in the aggregated corpus. In the construction of our dataset, we use the text (if any) of the Twitter posts and not the respective articles. Before doing so, we preprocess both the corpus of authoritative texts and

[2] Our data collection ended in December 2022. The platform is currently identified as "X". The list of newspapers included in our data set can be found in the appendix. The selection of newspapers was determined through a two-step process: first, we verified whether the newspapers had been included in prior studies (Baker et al. 2016) and were categorized as primary information outlets. Second, we expanded upon this criterion by evaluating whether each newspaper's Twitter handle had a substantial number of followers and maintained consistent posting activity.

Table 2.1: Dataset Summary

Country	Tot. no. News Outlets	Tot. no. Tweets	Tot. no. Words	Tot. no. Terms	Language
AR	4	1,892,464	16,592,918	50,513	Spanish
AU	4	539,792	5,805,411	46,253	English
BR	3	1,091,173	9,157,179	46,841	Portuguese
CA	5	1,071,360	9,662,873	48,695	English
CH	4	312,897	2,477,651	40,932	German and French
CL	3	1,088,455	9,422,589	48,409	Spanish
CN	3	488,076	6,786,626	53,342	English
CN (HK)	2	305,490	2,982,160	28,143	English
CO	3	1,291,773	10,348,974	55,984	Spanish
DE	4	708,692	6,779,556	31,190	German
ES	4	1,484,502	13,662,696	61,176	Spanish
FR	4	1,004,595	8,950,751	54,021	French
IN	4	1,909,474	22,370,252	71,505	English
IT	3	1,268,190	10,209,216	42,084	Italian
JP	4	353,247	3,819,349	36,544	English
KR	4	302,946	2,814,363	28,923	English
MX	4	2,041,505	19,553,684	70,880	Spanish
NO	3	152,521	921,258	13,022	Norwegian
NZ	3	147,218	1,156,784	20,538	English
PT	3	802,011	6,102,651	40,929	Portuguese
SA	3	914,623	8,754,709	44,571	Arabic
SE	4	190,493	1,607,852	19,201	Swedish
UK	4	791,929	6,978,125	46,460	English
US	11	2,735,338	29,141,511	72,662	English
ZA	3	453,135	3,945,579	33,473	English
Total	**96**	**23,341,899**	**220,004,717**	**1,106,291**	

Notes: This table shows summary statistics for all news that we collect for a large cross section of countries. Our real-time data range from 1/10/2014 to 31/12/2022. For each country, we report the total number of Twitter accounts that we monitor (major newspapers and other media outlets), the total number of tweets collected, total number of words, total number of unique terms (accounting for unigrams and bigrams), and original language. We translated the corpus in Engle et al. (2020) to 8 different languages, in order to construct our index.

the newspaper's tweets. The preprocess removes (1) any links, (2) any tags to Twitter accounts, (3) any special characters and/or numbers, and (4) stop words (e.g, "and" "or"). Furthermore, it stems each word (using the Snowball algorithm), and it keeps only words with more than two characters. The stemming procedure retains only the root of each terms; for example, terms like "work" and "working" are included as unique terms. See Figure A.1 in the appendix for a word cloud representation of the preprocessed corpus in English.

Second, we construct the analogous count from our preprocessed Twitter texts for every country. Our approach enables us to construct our measures both at high and low frequency. When studying the impact of climate attention shocks on currencies, we use a daily aggregation method. In what follows, we focus on monthly aggregation of tweets for illustrative purposes.

We refer to the Twitter texts aggregated in a specific month for a specific country as "document". At the monthly frequency, we have a total of 99 documents for each country. Next, we convert the term counts from each one of our documents into term frequency–inverse document frequency scores (in short, "tf-idf" scores). Finally, for each document, we compute the cosine similarity between its tf-idf scores and those of the entire corpus. We carry out the described procedure for every country within our sample.

In cases where a country uses a language other than English, we employ the Google Translate API to translate the original English corpus into the appropriate foreign language. These languages include German, Spanish, French, Norwegian, Swedish, Arabic, Portuguese, and Italian. This approach enables us to create Climate Change Vocabulary (CCV) sets for each of the available languages, facilitating the analysis of newspapers published in local languages. This approach mitigates the potential news bias that might arise from solely relying on English-language newspapers in non-English-speaking countries.[3] For countries with multiple local languages, such as Switzerland, we generate an index for each language and calculate the average across these indices.

Figure 2.1 presents our weekly Climate Attention Index (CAI), which has been aggregated across countries. The index is constructed by taking the equally weighted average across all the countries in our sample. The calculation of the index at the country-level is necessary because we cannot mix documents written in different languages. Consolidating documents at the country-language-frequency level would yield highly comparable results.

Our index effectively reflects specific global events. For instance, the United Nations Climate Change Conference (COP26), held in early November 2021 with the aim of promoting commitments to combat climate change, garnered significant attention worldwide, as evident in our index. Conversely, the COVID-19 pandemic and the invasion of Ukraine had the opposite effect on our index, with a notable decline attributed to a reduced proportion of climate-related discussions during these periods. Typically, significant events like elections, conflicts, and health crises tend to lead to a decrease in our index.

Figure 2.2 illustrates the same index across all of the countries included in our dataset. One of the novelties of our dataset is that it spans a large cross-section, which enables us to measure heterogeneity across countries. In the spirit of the international macro-finance literature, in what follows we focus primarily on heterogeneous exposures to global shocks (see Lustig et al. (2011)).

Specifically, let $CAI_{i,t}$ denote the index for country i at time t. The world climate attention index is defined as follows:

$$\overline{CAI}_t = \sum_i \omega_{i,t} CAI_{i,t}, \tag{2.1}$$

[3]For China, Japan and Korea, we retained the English-speaking newspapers posts due to the challenges in preprocessing languages with special characters. Our results are robust to the removal of these countries.

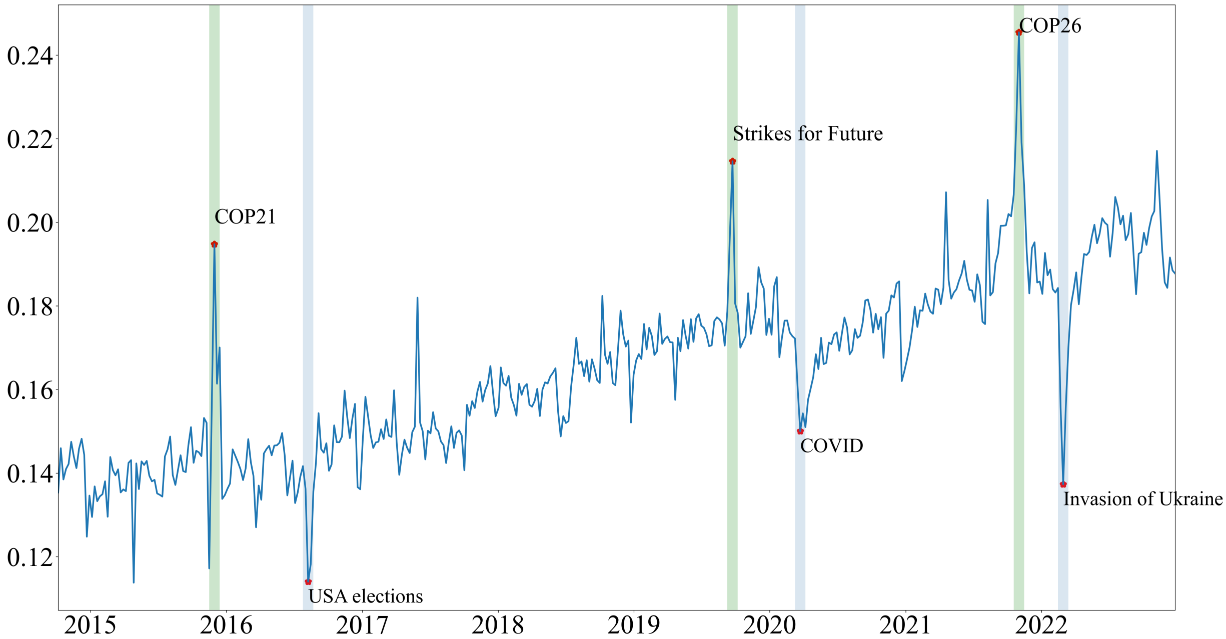


Figure 2.1: World Climate Attention Index (CAI)

Notes: This figure shows our Global Climate Attention Index. It is based on the methods described in section 2.2 applied to the countries listed in table 2.1 and a total of 8 languages.

where $\omega_{i,t}$ is a country-specific weight. We provide results based on three alternative weighting schemes: (i) by equally weighting all countries, (ii) by weighting each country proportionally to its share of global GDP, and (iii) by weighting each country proportionally to its share of Twitter volume. Volume weights are based only on tweets with a similarity score with the corpus of authoritative text above median. Equivalently, we only include the volume of tweets that are the most relevant for the focus of our analysis. We then estimate the following regression:

$$\Delta CAI_{i,t} = \beta_i \cdot \Delta\overline{CAI}_t + u_{i,t}, \tag{2.2}$$

where Δ refers to a demeaned growth rate over time for each country. We regard β_i as a measure of exposure to a common, i.e., global, component, and $u_{i,t}$ as a pure local component.

Table 2.2 presents the estimated β parameters for each country in our sample under various weighting schemes for the global CAI index. These β values exhibit a strong positive correlation across different weighting schemes, with correlation coefficients exceeding 0.9. Using the equally weighted global CAI index as the reference, we observe that South Korea, Saudi Arabia, China (including both Hong Kong and Mainland), and Japan consistently possess the highest β values. This aligns with expectations, as some of these countries are typically perceived as being more susceptible to climate change impacts (see for example, Hijioka et al. (2014), Pigato and Stewart (2020), and Sun et al. (2022)), reinforcing the credibility of our indices. Conversely, Switzerland, South Africa, Chile, Norway, and Germany consistently exhibit the lowest β values, indicating their relative resilience to climate change effects.

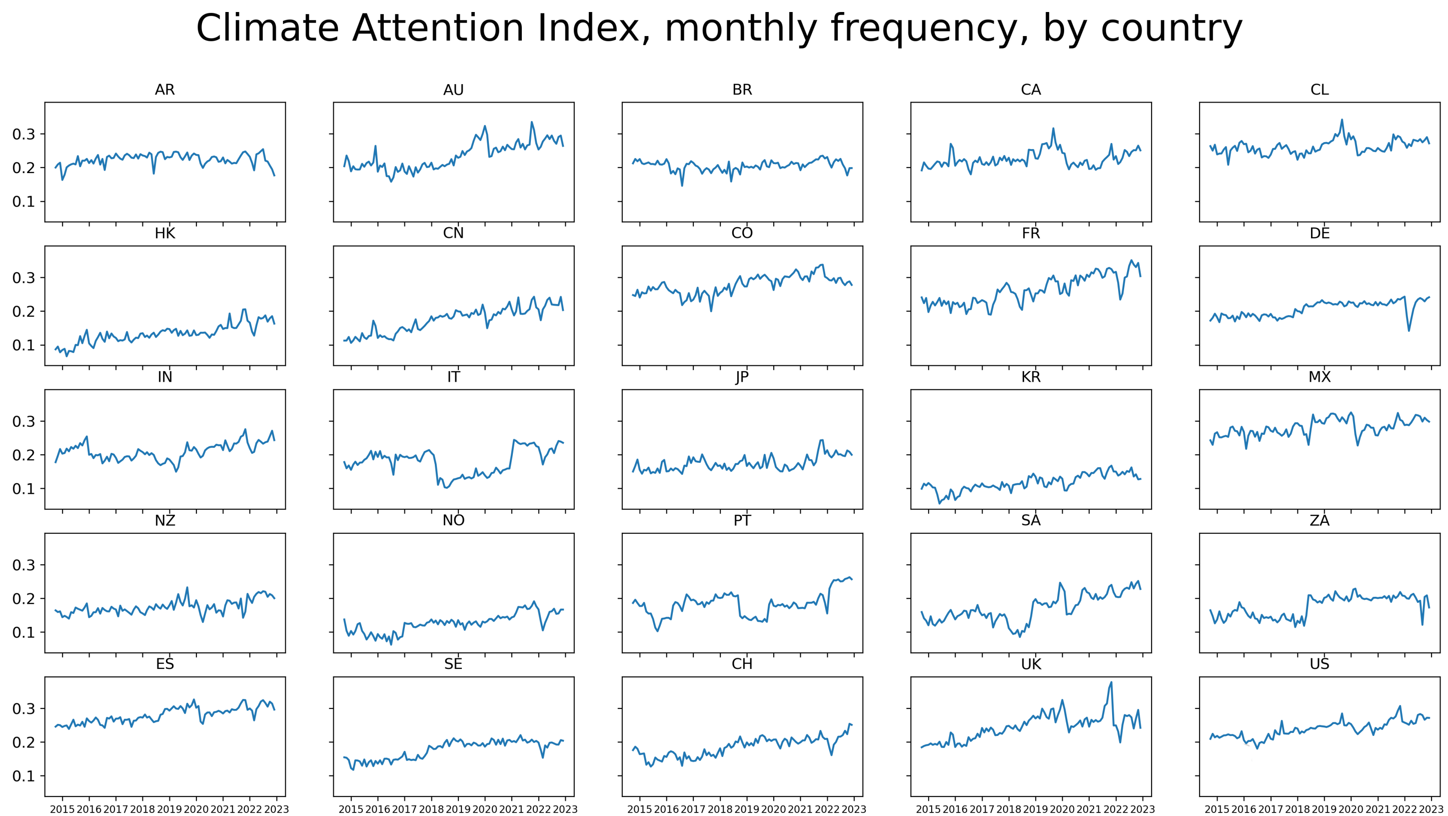


Figure 2.2: World Climate Attention Index (CAI) by country

Notes: This figure shows our Global Climate Attention Index. It is based on the methods described in section 2.2 applied to the countries listed in table 2.1 and 8 languages. For countries with multiple languages (e.g., Switzerland), the country-level index is the average of each language index.

Table 2.2: Estimated β for each country

	Equally weighted		GDP weighted		Volume weighted	
Country code	β	s.e.	β	s.e.	β	s.e.
AR	0.929	0.209	0.574	0.148	1.047	0.246
AU	1.419	0.197	0.932	0.190	1.375	0.219
BR	0.616	0.293	0.470	0.218	0.717	0.294
CA	1.710	0.224	1.312	0.138	1.789	0.232
CL	0.241	0.285	0.113	0.175	0.262	0.291
CN (HK)	2.105	0.325	1.196	0.326	1.619	0.518
CN	2.014	0.344	1.787	0.142	1.887	0.450
CO	0.679	0.180	0.561	0.115	0.702	0.207
FR	0.662	0.247	0.320	0.243	0.712	0.274
DE	0.363	0.122	0.236	0.074	0.316	0.160
IN	1.410	0.367	1.144	0.184	1.547	0.330
IT	0.985	0.593	0.624	0.473	1.441	0.705
JP	1.895	0.271	1.472	0.236	1.750	0.233
KR	2.413	0.394	1.624	0.280	1.889	0.324
MX	0.613	0.243	0.376	0.141	0.847	0.244
NZ	0.811	0.293	0.428	0.223	0.767	0.318
NO	0.344	0.320	0.245	0.267	0.166	0.313
PT	0.469	0.663	0.083	0.432	0.130	0.614
SA	2.266	0.574	1.108	0.251	2.254	0.609
ZA	0.197	0.466	-0.029	0.466	0.063	0.612
ES	0.534	0.128	0.366	0.084	0.578	0.114
SE	0.592	0.251	0.255	0.116	0.394	0.208
CH	-0.028	0.414	-0.014	0.263	-0.168	0.335
UK	0.993	0.269	0.798	0.214	0.998	0.339
US	0.768	0.143	0.743	0.083	0.905	0.134

Notes: This table reports the estimated β_i reported in equation (2.2), that is, the country-specific exposure to the global climate attention shock. The global climate attention shock is constructed as the equally weighted, GDP-weighted, and Twitter volume-weighted cross-sectional average at each point in time. Volume- and GDP-based weights are lagged by one period. The sample ranges from 2015:Q1 to 2022:Q4. Standard errors are HAC-adjusted.

Top days. In the next section, we use high-frequency data to identify the impact of climate news shocks on exchange rates and stock returns. To identify innovations in our global index, we initially calculate the daily growth rate in the raw CAI index for each country and subsequently remove its mean. These demeaned indices at the country level are then winsoried at the 99.9% and aggregated to generate a global metric.[4] The outcomes of this analysis are visualized in Figure 2.3.

[4]When we compute our index at the daily frequency, it is possible that in some days and for some countries we could have extreme growth rates because our index gets close to zero. We remove both the top- and bottom-0.1% realizations from our sample.

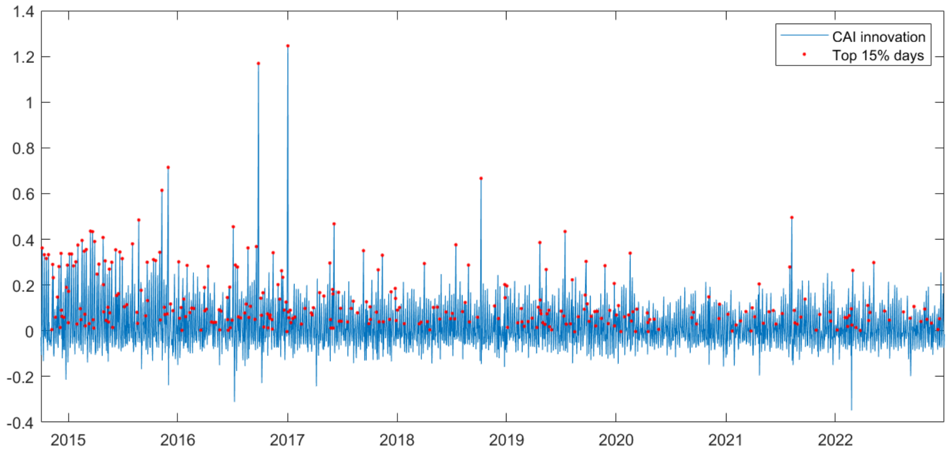


Figure 2.3: Innovations to World Climate Attention Index (CAI)

Notes: This figure shows innovations to our equal-weights Global Climate Attention Index constructed as described in section 2.2. Our cross section of countries is reported in table 2.1. Country-level indices are aggregated using equal weights. Red dots are classified as climate-related news days (top-15% realizations). Weekends are excluded.

Based on the distribution of innovations in our global climate index, we select the top 15% of these innovations.[5] Specifically, we examine the reaction of currencies and stocks in the days following a substantial increase in climate attention. This approach serves two purposes: (i) to minimize noise from news unrelated to climate, and (ii) to more effectively identify days driven by "pure" climate news shocks. Such a methodology is commonly used in the literature investigating the effects of news related to policy announcements (see, for instance, Leombroni et al. (2021)). It is worth noting that some days with a high volume of climate-related news may fall below our top 15% threshold due to the dominance of other topics in the news. As a result, similar to previous studies, this procedure may introduce a downward bias, making it less likely to detect a significant effect of climate news on asset prices and currencies.

Topics for top days. After identifying top-attention days, we examine the content of the news shared during these periods. Following the text analysis literature (see, for example, Hassan et al. (2019)), we employ a pre-trained sentence embedding model to extract semantic representations of newspaper tweets. These embeddings are then clustered to identify a smaller set of key themes, or main topics. This dimensionality reduction enables the use of topic-level cosine similarity scores to classify tweets as climate-related and further distinguish between those primarily associated with physical risk and transition risk. The process involves several carefully designed steps aimed at minimizing discretion on our part, which we detail below.

First, we generate 100 climate risk-related sentences using ChatGPT: 50 focused on physi-

[5]We analyze the distribution of news for each day of the week, which allows us to account for any patterns specific to particular days (e.g., Monday or Friday effects). This approach is standard in studies focusing on daily news.

cal risks (e.g., extreme weather) and 50 on transition risks (e.g., regulatory changes). The full list of these sentences is provided in Table A.2 in the Appendix. Next, we analyze the identified top climate-attention days by preprocessing raw newspaper tweets, removing non-informative elements such as links and special characters. We then convert the cleaned tweets into high-dimensional vector representations (sentence embeddings), which serve as inputs for the BERTopic clustering algorithm (see Grootendorst (2022)).[6]

After identifying the key topics discussed in the media on these top-attention days, we compute cosine similarities between each topic embedding and each AI-generated sentence embedding. This approach allows us to assess how closely each topic aligns with climate-related themes. Specifically, a topic is classified as climate-related if its cosine similarity score with any AI-generated climate sentence falls within the top 0.1% of the score distribution (with a threshold of 0.546). Using the 99.9^{th} percentile ensures a stringent definition of climate-related topics.

Finally, we classify each climate-related topic as either "Physical Risk" or "Transition Risk" based on its closest match to the sentences listed in Table A.2. If the most related sentence corresponds to physical (transition) risks, the topic is labeled accordingly. Figure 2.4 visualizes the most frequent words in these two groups of tweets, with the resulting word clouds reinforcing our methodology by clearly distinguishing between different sources of climate-related concerns. In the following analysis, we use these two groups of tweets in both our sentiment analysis and high-frequency asset pricing analysis.

Sentiment of tweets. We analyze the sentiment of climate-related tweets during top days by using the Twitter-XLM-RoBERTa model from CardiffNLP.[7] This is a transformer-based multilingual sentiment classifier suitable for text in multiple languages (Barbieri et al. 2022). Each one of our tweets is tokenized and passed through the pre-trained model in order to obtain a sentiment label (negative, neutral, or positive) based on the highest logit value.

Figure 2.5 presents the share of climate-related tweets in each risk category (physical and transition) that are classified as either positive or negative. In our sample, global news shocks related to climate typically have a negative tone. Equivalently, our top-attention days are usually bad-news days.

[6]BERTopic is a topic modeling technique that integrates transformer-based embeddings with clustering to extract coherent topics from large text datasets. We use the package provided at https://maartengr.github.io/BERTopic/index.html. We keep the following default options: (i) UMAP (Uniform Manifold Approximation and Projection) with five components, and (ii) HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) to group tweets into distinct topics. Since our data comprise tweets in different languages, we must use the embedding model "distiluse-base-multilingual-cased-v2" by Reimers and Gurevych (2019).

[7]Available at https://huggingface.co/cardiffnlp/twitter-xlm-roberta-base-sentiment.

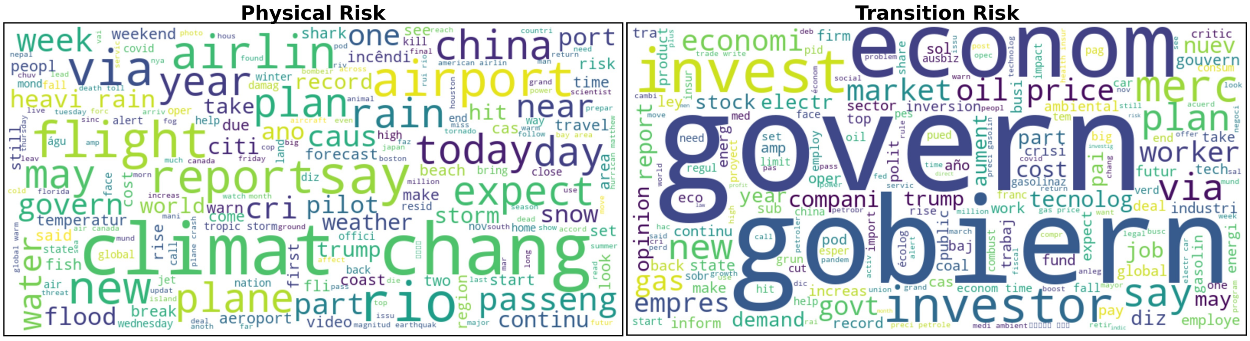


Figure 2.4: WordCloud of climate-related tweets during Top Days.

Notes: This figure shows the word cloud of root words extracted from climate-related Tweets in our sample of top days. Tweets are classified as climate-related if the cosine similarity of their associated topic exceeds the 99.9th percentile of the similarity scores calculated against a set of AI-generated reference sentences. Tweets are further classified as physical (transition) risk if their topic aligns most closely with a physical (transition) risk-related sentence (see table A.2). Our sample period starts in October 2014 and ends in December 2022. Transition risk generally refers to risks arising from changes in policies, technologies, and market preferences associated with climate action. Physical risk relates to the direct impact of climate change, such as extreme weather events and rising temperatures.

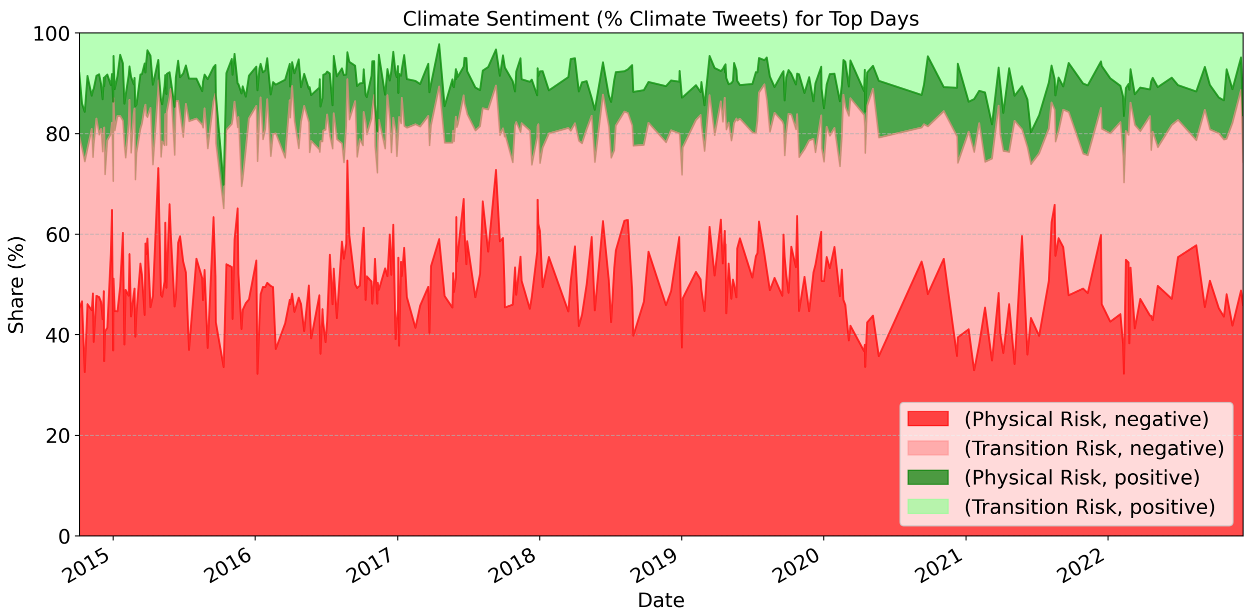


Figure 2.5: Sentiment of climate related tweets during Top Days.

Notes: This figure presents the share of climate-related tweets categorized by sentiment (positive or negative) and risk type (transition or physical) on days for which there is a positive spike (top-15%) in the global CAI innovations. The sentiment classification is based on a multilingual sentiment classifier Barbieri et al. 2022). The word clouds for transition and physical risk risk are reported in figure 2.4.

Attention vs vulnerability. To further analyze our text-based indices, we assess the correlation between our estimated β values and alternative measures of climate exposure. We first use the vulnerability score provided by the Notre Dame Global Adaptation Initiative. This index aims to encapsulate "a country's exposure, sensitivity, and capacity to adapt to the negative effects of climate change." We calculate the average of these scores over the longest available

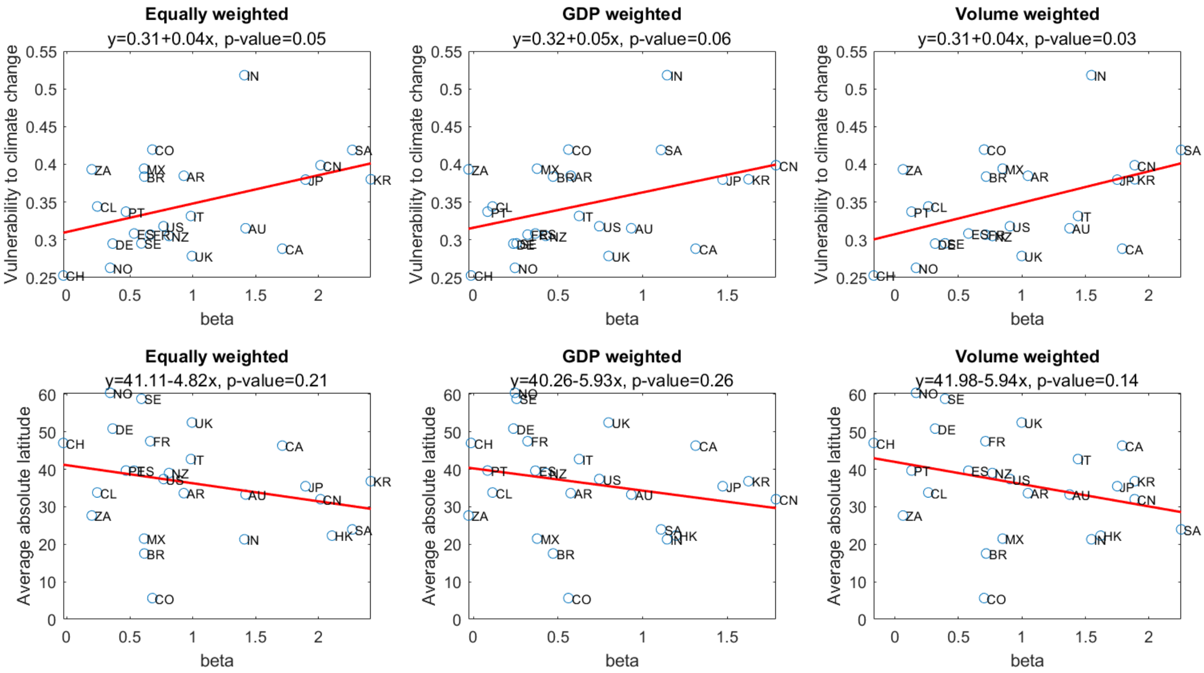


Figure 2.6: β against Vulnerability to Climate Change.

Notes: This figure shows the scatter plot of the estimated βs in equation (2.2), and measures of vulnerability to climate change across countries. The βs are estimated using the global climate attention shock constructed as either the equally-, GDP-, or Twitter volume-weighted cross-sectional average at each point in time. The upper panels consider each country's average vulnerability score from the Notre Dame Global Adaptation Initiative. The lower panels refer to the average absolute latitude of each country, that is, a population-weighted average absolute latitude of all cities in each country. City-level latitudes are collected from the World Cities Database.

period (1995-2021). The upper panel of Figure 2.6 illustrates the relationship between our β estimates and the vulnerability score. We observe a positive and statistically significant correlation between our β estimates and this index, implying that countries with greater susceptibility to climate change usually exhibit higher β values, i.e, more attention exposure to global climate news.

Second, we relate our estimated β values to absolute latitude. Generally, countries situated in regions with higher absolute latitudes are usually considered less exposed to rising temperatures resulting from climate change. To assess this, we compile a dataset of city-level absolute latitudes for each country and compute the population-weighted average of these latitudes. Our analysis reveals a negative correlation between our β estimates and absolute latitude, reinforcing the effectiveness of these estimates in capturing a country's exposure to climate change.

Importantly, we observe that both absolute latitude and the vulnerability index account for a moderate portion of the cross-sectional dispersion of our β coefficients. This indicates that our exposure measure is not fully captured by other existing indicators. In the next section, we will show that our exposure coefficients provide additional information beyond what is captured by vulnerability.

2.3 Climate news shocks in a no-arbitrage model

In this section, we use a no-arbitrage approach to derive a set of hypothesis that we test in our novel dataset.

2.3.1 Exchange Rates

Consider a complete market model with two countries, A and B, each populated by a representative investor. Assume that the stochastic discount factors in log-units are:

$$m_{A,t} = \bar{m} + \beta_A \cdot \epsilon_t, \quad m_{B,t} = \bar{m} + \beta_B \cdot \epsilon_t,$$

where positive realizations of ϵ_t represent common, i.e., global, bad news shocks about climate. If markets are complete, the percentage variation in the exchange rate can be expressed as follows:

$$\Delta e_{(A,B),t} = m_{A,t} - m_{B,t} = (\beta_A - \beta_B)\epsilon_t.$$

Here $\Delta e_{(A,B)} > 0$ refers to an appreciation of the currency of country A. In what follows, we assume that $\beta_A > \beta_B$, so that–by no arbitrage–adverse global climate news shocks should imply an appreciation of the currency of the most exposed country.

In order to test this condition, we obtain bilateral exchange rate data from Bloomberg and sort our country pairs as follows. Let A and B be two countries in a country-pair. For every non-repeated pairwise combination of countries, $i \neq j$, we look at the exposure of the country-specific CAI to the global CAI as estimated in equation (2.2). If $\beta_i - \beta_j > 0$ ($\beta_i - \beta_j < 0$) we denote country i as A and j as B (j as A and i as B) and focus on the exchange rate $e_{(A,B)}$.

We test this assumption at a high-frequency level, particularly on days when climate-related news is exceptionally significant. Figure 2.7 illustrates the average cumulative exchange rate variation following these specific days,

$$\Delta e_{t,t+k} := \sum_{s=0}^{k-1} \Delta e_{t+s,t+s+1} \frac{100 \cdot N^{top}}{k}.$$

Cumulative returns are averaged over the time horizon k, they are in percentage, and they are multiplied by N^{top}, i.e., the average number of top-days in a year. In our benchmark analysis $N^{top} = 38$.

To ensure that our observations are not influenced by idiosyncratic patterns in our dataset, these averages are computed relative to the average cumulative variation observed in the remaining 85% of business days in our sample. Specifically, we estimate the relative cumulative returns

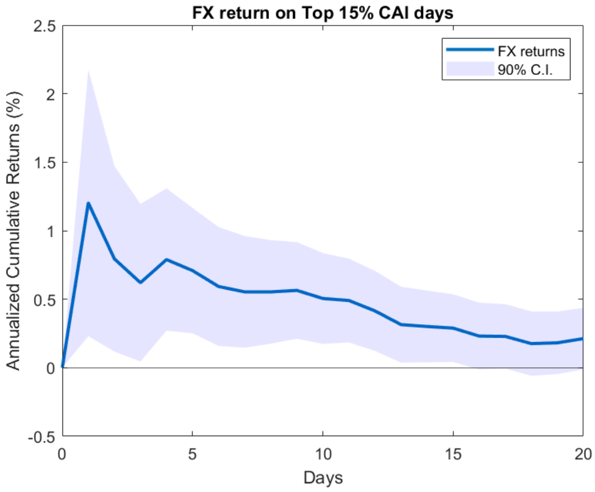


Figure 2.7: FX and Innovations to World Climate Attention Index (CAI)

Notes: This figure shows the average cumulative return in exchange rates around days with a positive spike (top-15% days in our sample) in the global CAI. This cumulative variation is net of the common cumulative variation that we obtain when considering all other days in our sample. Exchange rates are computed at Greenwich Mean Time in each day across 138 country pairs with exposures to the global CAI that are different from each other at the 10% confidence level. Exchange rates are quoted such that for each country pair an appreciation refers to the country with the higher exposure to global climate news shocks. Our Global Climate Attention Index is based on the methods described in section 2.2. Countries are equally weighted. The shaded areas present the 90% confidence interval which is constructed using standard errors clustered at both the date and country-pair level.

with the following regression:

$$\Delta e^{i,j}_{t,t+k} = a^{i,j}_k + b_k \cdot \not{\Bbbk}^{Top}_t + \epsilon^{i,j}_{t,k},$$

where $\not{\Bbbk}^{Top}_t$ is a dummy variable that equals 1 if the day falls in the set of extreme climate news days. b_k captures the relative cumulative returns over a horizon of k days. We only include country pairs where β_i is significantly larger than β_j at 10% level (see figure B.1, Appendix B). Standard errors are clustered at both the date and country-pair level.

If climate-related news has no bearing on exchange rates, we would anticipate null cumulative effects. Conversely, our findings reveal that adverse climate news shocks are linked to the contemporaneous appreciation of countries that are comparatively more exposed (i.e., possessing higher β values). Furthermore, this appreciation exhibits persistence and it extends over multiple days, signifying a lasting impact on currencies.

In Figure 2.8, we illustrate this effect among pairs of countries consisting solely of advanced economies (left panel), developing economies (right panels), or a combination of advanced and developing countries (middle panel).[8] The appreciation pattern is observed across all country pairs. Furthermore, it tends to persist longer when at least one emerging country is involved in the analysis.

[8]We define advanced economies according to the IMF, where the advanced economies in our sample are: AU, CA, CN (HK), FR, DE, IT, JP, KR, NZ, NO, PT, ES, SE, CH, UK, US. The developing economies are: AR, BR, CL, CN, CO, IN, MX, SA, ZA.

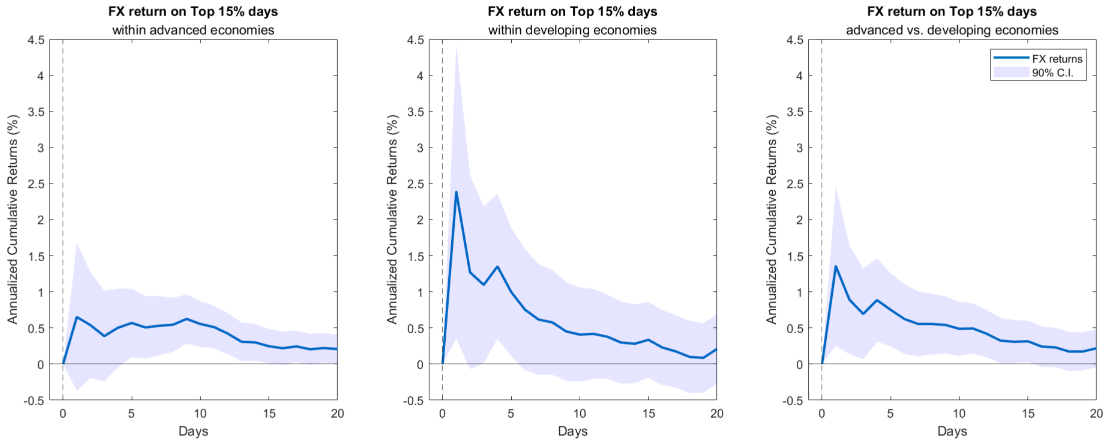


Figure 2.8: FX and Innovations to World CAI Across Country Pairs

Notes: This figure shows the average cumulative return in exchange rates after days with a top-15% positive spike in the global CAI. For more details, see the notes under figure 2.7 and section 2.2. We focus on country pairs in which (i) both countries are advanced economies (left panel), (ii) only one country is defined as an advanced economy (mid panel), and (iii) both countries are developing economies (right panel).

Robustness. In Appendix B, we show that our results are unchanged when we focus on either a GDP-weighted global index (figure B.2(a)) or a Twitter volume-weighted global index (figure B.2(b)). We also replicate the same exercise by focusing on GDP- and Volume-weighted CAI indeces (figures B.2(a)-B.2(b)). Our results are qualitatively unchanged. In addition, if we change the threshold of top days from 15% to either 10% or 20%, our results remains unchanged (figures B.3(a)-B.3(b)).

Figures B.4(a)-B.4(b) confirm that our results are not sensitive to the choice of the significance level that we use in order to select country-pairs with different exposure coefficients. Our results apply regardless of whether we include all country pairs (including those with modest heterogeneity in beta) or just those with a more pronounced degree of heterogeneity (5% confidence level).

Figure B.5(a) shows that our results are confirmed also when we aggregate the five Euro Area countries in our datatset (DE, FR, IT, PT, ES) into a single entity with a common currency. Most importantly, figure B.5(b) confirms that our results are not driven by vulnerability, but rather by the unique information content of our estimated exposure coefficients. Specifically, we orthogonilize our exposure coefficients with respect to vulnerability by considering the residuals of the cross sectional regression depicted in figure 2.6. The results in figure B.5(b) are based on the residual exposure and are consistent with our benchmark results. Equivalently, using heterogeneity in vulnerability across countries produces insignificant responses. The same conclusions apply to absolute latitude.

Figure B.6(a) confirms our main findings when using exposures estimated at the monthly frequency. Figure B.6(b) is constructed as figure 2.7 in main text, but it uses weekly observations and it includes weekends. The resulting weekly CAI index covers a total of 430 weeks. During this process, climate-related tweets are aggregated on a weekly basis, with each week of the time series commencing on a Saturday. This weekly aggregation provides a comprehensive perspective on climate-related discussions in social media, encompassing an extended period that includes weekends.[9] Similar to our daily frequency analysis, we identify weeks in which our index records top values and evaluate their cumulative impact on exchange rates.

In Figure B.7 in the appendix, we show the results from the following regression,

$$\Delta e^{i,j}_{t,t+k} = a^{i,j}_k + (b_k + c_k \cdot S_t) \cdot \mathbb{1}^{Top}_t + \epsilon^{i,j}_{t,k}, \tag{2.3}$$

where S_t denotes the share of tweets talking about transition risks at day t, normalized between 0 and 1. This specification enables us to identify whether our results are primarely driven by transition of physical risk. We find that the results depicted in figure 2.7 are consistent with those obtained by estimating equation (2.3) and setting S_t to its median value. Most importantly, we find that the results are stronger when $S_t = 1$, i.e., in top days when attention is entirely about transition risk. Thus transition risk seems to be the main driver of currency reactions.

In Figure B.8, we examine local top-attention days, as opposed to global top-attention days. We find that local shocks have no significant impact on FX markets. More broadly, local shocks also appear less relevant in the equity analysis presented next. For brevity, we exclude local shocks from the subsequent analysis.

2.3.2 Equity returns

Based on our findings related to currencies, we investigate the effects of our climate news index on equity returns denoted in USD. Specifically, we examine the behavior of average cumulative returns following days marked by a positive spike in our global index (CAI), taking into account both stock-specific exposure and country-specific conditions.

We use emission intensity as a proxy for stock-specific exposure, defined as emissions in tons of CO2 divided by sales. We source annual data on emission intensity from the Trucost dataset. For country-specific conditions, we use sensitivity coefficients denoted by β, as reported in table 2.2. Our dataset includes 16,774 firms and totals 23,305,563 firm-day observations. We estimate the following regression:

$$\begin{aligned} r^{i,c}_{t,t+k} &= \alpha^i_k + \zeta^{i,c}_{t,k} \cdot \mathbb{1}^{Top}_t + \epsilon^{i,c}_{t,k}, & (2.4) \\ \zeta^{i,c}_{t,k} &= \gamma_{0,k} + (\gamma_{1,k} + \gamma_{2,k} \cdot \beta_c) \cdot E_{i,t_a}, & (2.5) \end{aligned}$$

[9]To identify the top weeks in terms of climate attention, we calculate the growth rate of the weekly global CAI, then the top 10% of observations of this time series are considered as indicative of heightened climate attention. We only keep top weeks that are at least 5 weeks apart to avoid the confounding effects.

where c denotes the country in which the headquarter of the firm is located, i refers to one of our firms, β_c represents our equally weighted sensitivity, E_{i,t_a} is the emission intensity of firm i assessed the year prior to day t, and $\mathbb{K}_t^{Top}$ is a dummy variable that equals 1 if the day falls in the set of extreme climate news days. In equation (2.4), the left-hand side variable is defined as follows:

$$r_{t,t+k}^{i,c} = \sum_{s=0}^{k-1} r_{t+s,t+s+1}^{i,c} \frac{100 \cdot N^{top}}{k},$$

that is, it is averaged over the time horizon k, it is in percentage, and it is multiplied by the average number of top-days in a year.

Since emission intensity is very persistent at the annual frequency, the coefficient $\zeta_{t,k}^{i,c}$ is almost time-invariant in our dataset. Consequently, this composite coefficient primarily reflects heterogeneity across different countries and stocks. The literature has proxied the exposure of a firm to climate risk either by examining its emissions or considering geographical factors such as distance from the ocean or latitude. We hypothesize that a firm's equity returns' sensitivity to climate news reflects a convolution of its emission levels (as measured by total emissions) and the degree of investor attention to climate news within a specific country (as captured by our country-level β).[10] We are particularly interested in the interplay of these dimensions, as negative global climate news could be concentrated only among firms with high emissions in highly exposed countries ($\gamma_{2,k} < 0$). Our estimates confirm this intuition. In figure 2.9, we depict the estimates of our composite parameter $\hat{\zeta}_{t,k}^{i,c}$ over different daily horizons, k.

In the top portion of the figure, we first sort countries according to their exposure β. The top-decile of the distribution of βs is 2.27 (Saudi Arabia). The bottom-decile is 0.197 (South Africa). We then select the values of carbon emission intensity for a representative green (brown) firm using the 5th (95th) percentile of the emission intensity distribution, which corresponds to the median value within the top-decile (bottom-decile).[11] Given these numbers, we can depict the implied estimated composite coefficient $\zeta_{t,k}^{i,c}$ and its associated standard error across different daily horizons, k.

The bottom two panels convey the same information as the top panels but from a different perspective. Specifically, the bottom left (right) panel presents the composite coefficients for a representative green (brown) firm in a low-β country and in a high-β country.

Our analysis yields several significant findings. Firstly, firms situated in countries with moderate exposure to our global CAI tend to experience a more moderate loss of value. This effect holds for both high- and low-emission firms. Conversely, firms in countries with high β values

[10]This approach amounts to conjecturing that a brown firm in a country subject to floods may be subject to stronger valuation effects than an equally brown firm in a country whose climate is expected to improve due to global warming.

[11]The 5th (95th) percentile of the log emission intensity in our sample is 0.43 (6.88). The thresholds remain consistent both across and within countries because the country-level distribution of emission intensities is stable. See figure B.9 in the appendix for more details.

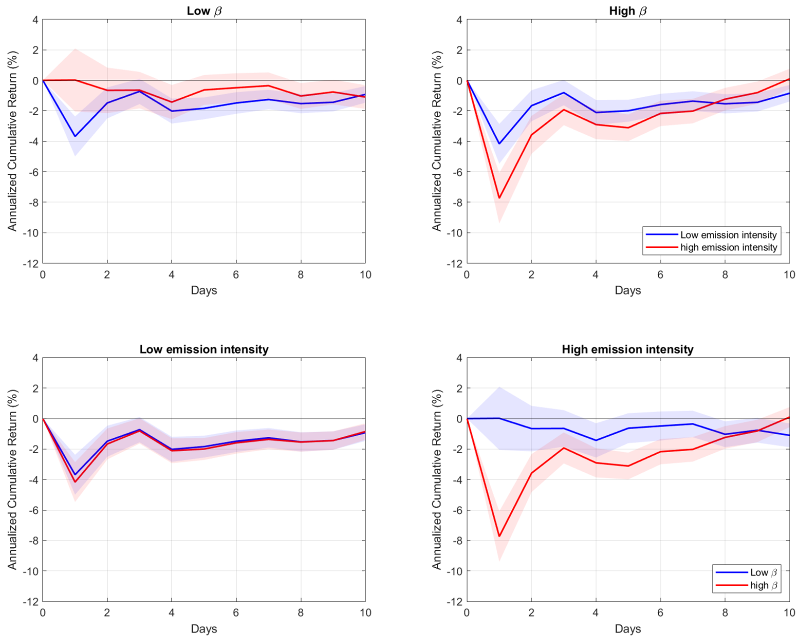


Figure 2.9: USD Equity Returns and Innovations to World CAI

Notes: This figure shows the average cumulative return in equity returns after days with a top-15% positive spike in the global CAI. For more details, see the notes under figure 2.7. We depict the estimates of $\zeta_{t,k}^{i,c}$ as defined in equations (2.4)-(2.5) across different daily horizons, k. Standard errors are clustered at the industry$\times$date and firm level. We follow the Fama-French 49 classification. The sample comprises a total of about 23 million firm-day observations. Daily returns are annualizedusing the average number of top climate days in a year. In the four panels, high (low) β refers to the top-decile (bottom-decile) of betas across countries in our sample, and high (low) emission intensity is using the top- and bottom-5% of emission-sorted firms in our sample. The shaded areas present the 90% confidence interval.

suffer a more severe loss of value, an effect that is particularly pronounced for high-emission firms. Across all these scenarios, the negative impact on cumulative returns is very persistent.

Changing the order of the sorting variables does not alter our conclusions. Specifically, the bottom two panels of figure 2.9 confirm that firms located in 'naturally hedged' countries (that is, countries with low β) are generally less affected by adverse climate news. However, when examining stocks with high emission intensity, the negative impact on valuation is more severe for firms in countries with a higher climate news β.

Following our analysis of currency markets, we examine whether these results are primarily driven by top-attention days in which news focuses on either physical or transition risk. Specifically, we estimate the following specification:

$$\begin{aligned} r_{t,t+k}^{i,c} &= \alpha_k^i + \zeta_{t,k}^{i,c} \cdot 1_t^{Top} + \epsilon_{t,k}^{i,c}, \\ \zeta_{t,k}^{i,c} &= (\gamma_{0,k} + (\gamma_{1,k} + \gamma_{2,k} \cdot \beta_c) \cdot E_{i,t_a}) + (\kappa_{0,k} + (\kappa_{1,k} + \kappa_{2,k} \cdot \beta_c) \cdot E_{i,t_a}) \cdot S_t, \end{aligned}$$

where S_t represents the share of tweets discussing transition risks on day t, normalized be-

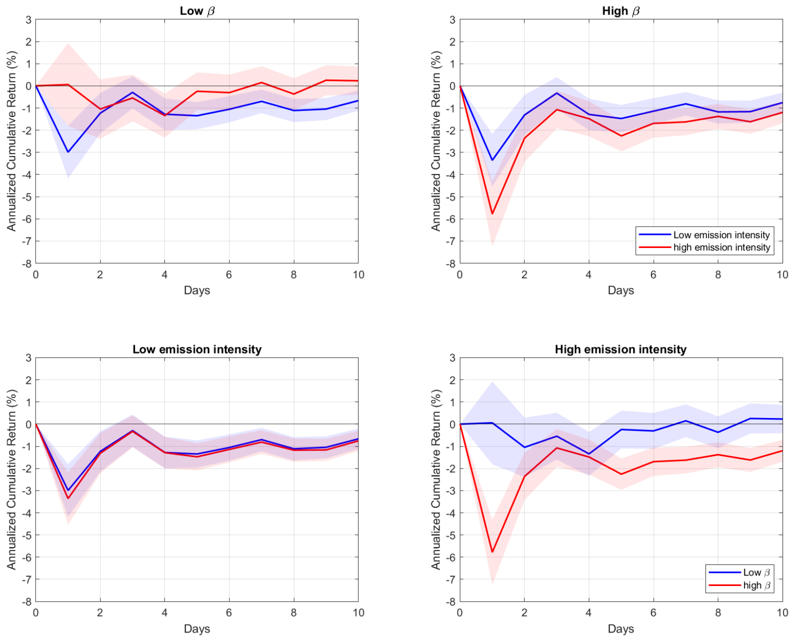


Figure 2.10: Local Equity Returns and Innovations to World CAI

Notes: This figure shows the average cumulative return in equity returns after days with a top-15% positive spike in the global CAI. Returns are in local units (as opposed to USD) and in percentage. For more details, see the notes under figure 2.9.

tween 0 and 1. Figure B.13 in the appendix presents our results. The responses are constructed using the composite coefficient $\zeta_{t,k}^{i,c}$, with $S_t = 0.50$ in the left panel and $S_t = 1$ in the right panel. For a median top-attention day, this specification reproduces the results reported in Figure 2.9. When news is primarily about transition risk, we observe more pronounced depreciations across all panels. As with currency markets, equity depreciations are largely driven by negative news on transition risk. Notably, the response of equity returns to these news events appears homogeneous across countries with varying exposure and across firms with different emissions. This finding suggests that (i) regulation-related interventions may have broadly similar effects across firms, and (ii) our benchmark result—where brown firms in more exposed countries experience greater depreciation—is primarily driven by top-attention days dominated by concerns over physical risk.

Most importantly, our results are not solely driven by the exchange rate adjustment. As shown in figure 2.10, our results hold also when we focus on returns expressed in local currency.

Robustness. In Appendix B, we show that our results are robust across various methodological changes. Panel (a) in Figure B.10 uses GDP-weighted CAI, while Panel (b) employs Twitter Volume-weighted CAI. Both panels confirm the robustness of our findings as documented in Figure 2.9. Furthermore, Figure B.11 compares the response of equity returns using thresholds of 10% and 20% to identify spikes in CAI, respectively. This figure documents that our results remain qualitatively unchanged regardless of the chosen threshold.

Figure B.12 provides supplementary analyses to validate our findings. In Panel (a) we cluster standard errors at the country$\times$industry$\times$date and firm levels, while in Panel (b) we use residual βs obtained by regressing our exposure coefficients on the University of Notre Dame's vulnerability index. These results corroborate our main conclusions.

2.4 Climate news shocks in general equilibrium.

The environment. The main equilibrium conditions of our model are derived in Appendix C. In what follows, we describe our economic environment and then derive the key predictions regarding trade and currencies. The model features two countries, home and foreign. Foreign variables are denoted by "*". Due to the symmetry of our environment, many of our equations are presented for the home country, with the implicit understanding that they apply to the foreign country as well. Time is discrete and we assume that there are only two periods denoted by the following three dates $t = 0, 1, 2$.

All shocks materialize at the end of the first period, i.e., at $t = 1$. Hence, at time $t = 0$ choices are made under uncertainty. At time $t = 2$, no additional shock materializes, or, equivalently, in the second period there is full resolution of uncertainty. Most importantly, all news received at time $t = 1$ affect the production frontier only at $t = 2$. Hence, from a time-1 perspective, these shocks represent pure news shocks.

In each country, the representative household has the following recursive preferences:

$$u_t = \begin{cases} (1-\beta)\log\left(C_t D_t^{-a}\right) + \frac{\beta}{1-\gamma}\log E_t\left[\exp\left\{u_{t+1}(1-\gamma)\right\}\right] & \gamma \neq 1 \\ (1-\beta)\log\left(C_t D_t^{-a}\right) + \beta E_t\left[u_{t+1}\right] & \gamma = 1 \end{cases}$$

When $\gamma = 1$, our preferences boil down to the case of log-expected utility. When $\gamma > 1$, instead, the households feature a preference for early resolution of uncertainty. Equivalently, households price news shocks, that is, their marginal utility immediately adjust with the arrival of news. Since we assume that there is full resolution of uncertainty at time $t = 1$, households are concerned about uncertainty about utility at time 1.

As in prior work, we assume that each agent consumes a bundle of two consumption goods:

$$C_t = X_t^{\lambda} Y_t^{(1-\lambda)}, \text{ and } \quad C_t^* = X_t^{*(1-\lambda)} Y_t^{*\lambda}, \quad t = 1, 2.$$

In what follows, we assume that X_t is the home-produced good, whereas Y_t is the foreign-produced good. The parameter $\lambda > 1/2$ is set so that there is home-bias in consumption. C_0 and C_0^* are both fixed paramters in this model.

D_t captures the damage caused by climate change and $0 < a < 1$ determines its impact on the household's utility. C_t denotes the consumption bundle at time t. We think of D_t as resulting from economic activity. Since in our model we abstract away from labor and capital is one-period ahead pre-determined, output at time 1 is pre-determined. As a result, D_0, D_0^*, D_1, and D_1^* are given in this setting. In contrast, at time $t = 2$, damages are endogenous and are modelled as follows:

$$D_2 = e^{bg} \left(e^{\theta} G/I_{g,1}\right)^{\lambda_g} \left(e^{-\theta} G^*/I_{g,1}^*\right)^{1-\lambda_g}, \quad D_2^* = e^{g} \left(e^{-\theta} G^*/I_{g,1}^*\right)^{\lambda_g} \left(e^{\theta} G/I_{g,1}\right)^{1-\lambda_g}.$$

The term e^g captures a global climate news shock with $g \sim N(0, \sigma_g^2)$. The exposure of the foreign country to this shock is normalized to 1, whereas b captures the exposure of the home country. When $b > 1$, the home country is more exposed than the foreign one. Therefore, $b \neq 1$ captures heterogenous exposure to a common climate shock. The parameter λ_g determines the relevance of home-originated emissions relative to foreign-originated emissions. When $\lambda_g = 0.5$, agents are equally concerned about domestic and foreign emissions. When $\lambda_g > 0.5$, local emissions produce stronger damages.

At time $t = 2$, output is assumed to be equal to $e^{\theta} G$ and $e^{-\theta} G^*$ in the home and foreign country, respectively. The term e^{θ} determines a local productivity shock, with $\theta \sim i.i.d.N(0, \sigma_\theta^2)$. When $\theta > 0$, the home country is $2\theta\%$ more productive than the foreign country. G and G^* are bundles of time-1 local and foreign investments in the home- and foreign-country, respectively:

$$G = I_{x,1}^{\lambda_I} I_{x,1}^{*1-\lambda_I}, \text{ and } G^* = I_{y,1}^{1-\lambda_I} I_{y,1}^{*\lambda_I},$$

Note that when $\lambda_I = \lambda$, the consumption bundle is identical to the investment bundle in each country. Given these assumptions, the damage in each country is a combination of economic activity (output) both at home and abroad.

Finally, $I_{g,1}$ and $I_{g,1}^*$ represent green investments in the home and foreign country, respectively. These components capture the ability of reducing the impact of output on climate by investing in new greener technologies.

At time $t = 1$, output is predetermined and fixed to 1 for simplicity. The implied resource constraints in the economy are standard,

$$1 = X_1 + X_1^* + I_{x,1} + I_{y,1} + I_{g,1}, \quad 1 = Y_1 + Y_1^* + I_{y,1}^* + I_{x,1}^* + I_{g,1}^*,$$

and account for both international flows of consumption and investment goods. At time $t = 2$, there is no incentive to further invest in capital as we are focusing on a 2-period model. Hence time-2 output is solely used to support international consumption:

$$e^{\theta} G = X_2 + X_2^*, \quad e^{-\theta} G^* = Y_2 + Y_2^*.$$

Returns. In this economy, we define the returns to both brown and green capital across periods. Between time-1 and time-2, all returns equal the risk-free rate because there is full resolution of uncertainty at time-1. Between time-0 and time-1, however, returns are random as they depend on stochastic fundamentals.

Let $r_{x,1} - r_{g,1}$ represent the excess return of a zero-dollar investment strategy that is long in home country brown capital and short in its green capital. In the Appendix, we demonstrate that the following equation holds:

$$r_{x,1} - r_{g,1} = constant + \lambda^{s}_{r_x - r_b}\lambda^{\theta}_{s}\theta + \lambda^{s}_{r_x - r_b}\lambda^{g}_{s}g, \tag{2.6}$$

where the terms denoted by λ^*_* are composite exposure coefficients whose signs can be analytically determined. In Appendix C, we address the planner's problem associated with this setting. Additionally, we explore an environment where emission externalities have been corrected and the economy is at its first-best. In this setting, we derive several testable hypotheses that we describe below.

Additional Testable hypotheses. In Appendix C, we prove the following two propositions.

Proposition 2.4.1. If $b > 1$, the brown-minus-green investment return in the home country depreciates when there is a positive global climate news shock g. In addition, the effect is stronger with a higher b.

In the appendix, we show that the coefficient $\lambda^{s}_{r_x - r_b} < 0$, and $\lambda^{g}_{s} > 0$ when b is greater than 1. This means that in countries with relatively high exposure, climate shocks have a greater impact on brown stocks. Additionally, $\lambda^{g}_{s} > 0$ increases with the level of country exposure, denoted by b. This finding aligns with our empirical approach where we consider the interaction between firm-level emissions and country exposure (equations (2.4) and (2.5)).

Proposition 2.4.2. Let $\gamma > 1$, when $b > 1$ ($b < 1$), a global climate news shock $g > 0$ causes an appreciation of home (foreign) real exchange rate.

Proposition 2 confirms what we have derived in the previous section when adopting a no-arbitrage approach. Namely, adverse global news shocks to climate cause an appreciation of currencies associated to countries relatively more exposed to these news.

Proposition 2.4.3. When $b > 1$, a global climate news shock $g > 0$ increases the relative pseudo Pareto weight of the home country, S_1, and it decreases the net exports of consumption and investment goods, $\frac{NX^C_1}{X_1}$ and $\frac{NX^I_1}{I_x}$.

In order to better explain proposition 3, we find it appropriate to focus on the equilibrium the net exports of consumption goods,

$$\frac{NX_1^C}{X_1} - \frac{NX_1^{C*}}{Y_1^*} = \beta_\theta^{NX} \cdot 2\theta + \beta_g^{NX}(b-1)g.$$

In the appendix, we prove that under the assumptions detailed in our propositions the coefficients in front of the three fundamental shocks have the following structure and signs:

$$\begin{aligned}
\beta_\theta^{NX} &= -\frac{1}{2}\frac{1-\lambda}{\lambda}(e^{\bar{s}_1} - e^{-\bar{s}_1})\lambda_s^\theta > 0 \\
\beta_g^{NX} &= -\frac{1}{2}\frac{1-\lambda}{\lambda}(e^{\bar{s}_1} - e^{-\bar{s}_1})\frac{\lambda_s^g}{b-1} < 0.
\end{aligned}$$

Since similar results hold for net exports of investment goods, the sign of these coefficients apply also to the difference of the total net exports. For the home country, a positive productivity shock generates an outflow of resources to the foreign country. In contrast, an adverse shock to global climate generates an outflow only if the home country has a moderate exposure to it, $b < 1$.

Empirical tests for international trade. In the data, we are interested in estimating the following system of equations:

$$\Delta CAI_{i,t} = \beta_i \cdot \overline{\Delta CAI}_t + u_{i,t} \tag{2.7}$$

$$\begin{aligned}
\Delta\left(\frac{NX_{i,t}}{GDP_{i,t}}\right) - \Delta\left(\frac{NX_{j,t}}{GDP_{j,t}}\right) &= \Gamma \cdot (\beta_i - \beta_j)\,\overline{\Delta CAI}_t + ... \\
&\quad + \theta' \cdot (control_{i,t} - control_{j,t}) + \epsilon_{ij,t},
\end{aligned} \tag{2.8}$$

where $\frac{NX_{i,t}}{GDP_{i,t}}$ is the net export over GDP of country i at time t. We add the first difference to make sure that all variables are stationary. $\Delta CAI_{i,t}$ denotes the growth rate of the climate attention index for country i and quarter t, $\overline{\Delta CAI}_t$ is the cross-sectional average of ΔCAI_t at date t, i.e., the global CAI index.[12] The term $control$ is a vector of control variables including the change in the industrial production index (ΔIPI), to control for country-specific productivity shocks, and the share of Twitter volume on days when the CAI is at the bottom 5%, to control for other events that affect Twitter activities. In the appendix, we consider a more general version of our model which features also local climate news shocks with volatility σ_z. In main text, we abstract away from local shocks by setting $\sigma_z = 0$ because in our international setting global shocks feature a prominent role.

We provide results based on the cases in which $\overline{\Delta CAI}_t$ is constructed in one of three ways:

[12] In order to eliminate the effect of the Covid-19 and the beginning of the war in Ukraine, we have regressed $\Delta CAI_{i,t}$ on two dummies of 2020Q2 and 2022Q1 for each country, and obtain the residual for our analysis.

(i) by equally weighting all countries, (ii) by weighting each country proportionally to its share of global GDP, and (iii) by weighting each country proportionally to its share of Twitter volumes.

Simple OLS on both equations will produce incorrect inference for the coefficient Γ, because the β coefficients in equation (2.8) are themselves estimated from equation (2.7). We address this econometric issue by estimating the system via GMM with a pre-determined weighting matrix. Specifically, the set of moment conditions is

$$g_T(\beta, \gamma, \theta) = E_T \left[\begin{array}{c} u_{i,t}\overline{\Delta CAI}_t \\ \epsilon_{ij,t}\left(\beta_i - \beta_j\right)\overline{\Delta CAI}_t \\ \epsilon_{ij,t}\left(control_{i,t} - control_{j,t}\right) \end{array} \right] = E_T \left[\begin{array}{c} g_1(\beta) \\ g_2(\beta, \Gamma, \theta) \end{array} \right]$$

where $g_1(\beta)$ is the vector of orthogonality conditions associated to equation (2.7) for each country i, and $g_2(\beta, \Gamma, \theta)$ stacks the three orthogonality conditions associated to equation (2.8) for each country pair $\{i, j\}$. The first block g_1 pins down β, and the second block g_2 pins down the remaining coefficients Γ, θ. Given two control variables, if we have N countries in our sample, the number of moments will be $N + 2N \times (N - 1)$, and the number of parameters is $N + 4$. This is an over-identified GMM problem for $N > 2$.

Next, we follow Cochrane (2009) and estimate the following linear combination of the moment conditions:

$$a_T g_T = 0$$

where

$$a_T = \left[\begin{array}{cc} I_N & 0 \\ 0 & \frac{\partial g_2(\beta,\gamma,\theta)}{\partial[\gamma,\theta]} W \end{array} \right]$$

with I_N and W being identity matrices of size N and $2N \times (N - 1)$, respectively. Weighting the block of orthogonality conditions of equation 2.7 with an identity matrix ensures that the estimated vector of βs coincides with OLS estimates. The set of moments conditions in g_2 is instead efficiently estimated. Denote by $b = [\beta, \gamma, \theta]'$ the vector of all parameters. The asymptotic distribution of b is given by

$$\sqrt{T}(\hat{b} - b) \sim N\left(0, (ad)^{-1} aSa'(ad)'^{-1}\right)$$

where $a = \lim_{T\to\infty} a_T$, $d = \frac{\partial g_t}{\partial b}$ is the gradient, and S is the spectral density matrix estimated using the residuals.

Given our interest on the role of global climate news shocks on the dynamics of international trade, we focus on the estimated parameter Γ in equation (2.8). According to our model, this coefficient should be negative. We estimate the parameters of interest using an expanding window to show whether the estimated coefficient is stable over different sample periods. Specifically,

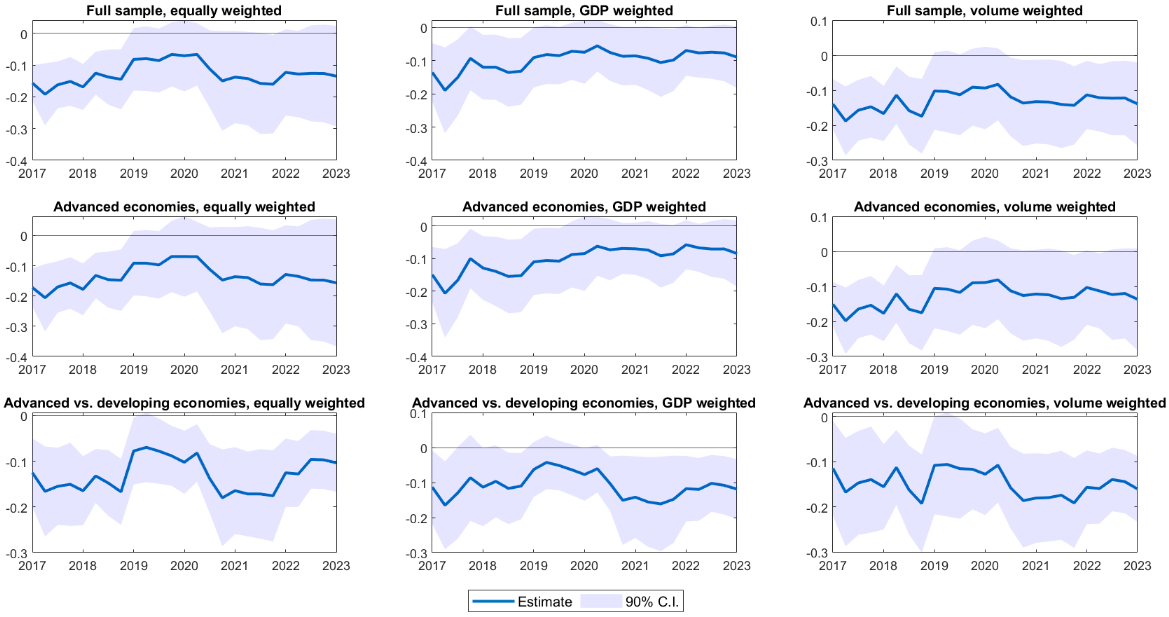


Figure 2.11: GMM estimation of Γ.

Notes: The figure shows estimates of Γ (defined in equation (2.8)) using expanding time windows. Each window starts in 2015:Q1 and ends on the date reported on the x-axis. The parameter Γ is estimated using the GMM described in section 2.4. Panels in different columns refer to results obtained from different ways to aggregate country-level CAI indices in order to form a global component. Panels in different rows refer to results obtained from different groups of country-pairs. Advanced economies are defined according to the IMF. Standard errors are HAC-adjusted. The shaded areas present the 90% confidence interval.

we fix our starting period at 2015:Q1 (where our sample begins), and expand the end of sample from 2016:Q4 to 2022:Q4. The estimated Γ coefficient along with its 90% confidence interval are presented in figure 2.11.

Our estimates support the prediction of our model, as they suggest that an adverse global climate shock produces an outflow of resources away from countries with a lower exposure. This result is confirmed across different ways to aggregate our country-level CAI in order to form a global index. In addition, our result is driven mainly by country pairs in which there is at least one developed economy. This is broadly consistent with the spirit of our model since we assume that risk sharing is implemented with frictionless trade.

2.5 Concluding Remarks

Our research significantly contributes to our understanding of the impact of global climate attention on economic and financial outcomes. The availability of high frequency data is particularly valuable as it enables a more direct analysis of how climate news shocks reverberate throughout global financial markets and influence currencies on a global scale. The establishment of a clear connection between climate attention shocks and bilateral trade among nations represents a noteworthy advancement for both the macro-finance literature and policymakers.

Moreover, the economic model we introduce, along with the climate attention index we propose, open up various avenues for future research. For instance, our economic theory implies that heightened global news attention to climate issues may influence decisions regarding investments in green technology. Furthermore, the climate attention index we have developed can be leveraged to explore the relationship between climate-related news and returns on a wide range of assets above and beyond stocks, both within individual countries and across international borders. These potential research directions hold promise for gaining deeper insights into the complex interplay between climate awareness and economic and financial dynamics.

Establishing a connection between attention to climate news and currency movements is of paramount importance in the current global landscape. As climate change becomes an increasingly pressing concern, understanding how shifts in climate attention impact financial markets and trade can inform both economic policies and investment strategies. It provides a valuable tool for policymakers to anticipate and respond to economic implications of climate-related events and offers investors insights into the potential risks and opportunities associated with climate awareness. By shedding light on these intricate connections, our research serves not only the academic community but also contributes to the broader efforts to address the multifaceted challenges posed by climate change on a global scale.

Bibliography

Acharya, V. V., T. Johnson, S. Sundaresan, and T. Tomunen. Is physical climate risk priced? evidence from regional variation in exposure to heat stress. Technical Report w30445 National Bureau of Economic Research 2022 https://www.nber.org/papers/w30445.

Baker, S. R., N. Bloom, and S. J. Davis. 2016. Measuring economic policy uncertainty*. *The Quarterly Journal of Economics* 131(4):1593.

Bansal, R., D. Kiku, and M. Ochoa. Climate change risk. Technical report Federal Reserve Bank of San Francisco 2019 https://www.frbsf.org/economic-research/events/2019/november/economics-of-climate-change/files/Paper-5-2019-11-8-Kiku-1PM-1st-paper.pdf.

Barbieri, F., L. Espinosa Anke, and J. Camacho-Collados. XLM-T: Multilingual language models in Twitter for sentiment analysis and beyond. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. European Language Resources Association 06 2022.

Bolton, P., and M. Kacperczyk. 2021. Do investors care about carbon risk? *Journal of Financial Economics* 142(2):517–549. doi: 10.1016/j.jfineco.2021.05.008.

Bolton, P., and M. Kacperczyk. 2023. Global pricing of carbon-transition risk. *Journal of Finance* 78 (6):3677–3754. doi: 10.1111/jofi.13272.

Bolton, P., M. T. Kacperczyk, and M. Wiedemann. The co2 question: Technical progress and the climate crisis. Technical report Working Paper 2023 https://ssrn.com/abstract=4212567.

Cochrane, J. *Asset pricing: Revised edition*. Princeton university press 2009.

Colacito, R., M. Croce, S. Ho, and P. Howard. November 2018. Bkk the ez way: International long-run growth news and capital flows. *American Economic Review* 108(11):3416–49.

Colacito, R., B. Hoffmann, and T. Phan. 2019. Temperature and growth: A panel analysis of the united states. *Journal of Money, Credit and Banking* 51(2-3):313–368. doi: https://doi.org/10.1111/jmcb.12574.

Della Corte, P., L. Sarno, and I. Tsiakas. 2009. An Economic Evaluation of Empirical Exchange Rate Models. *Review of Financial Studies* 22(9):2491–530.

Della Corte, P., L. Sarno, and I. Tsiakas. 2011. Spot and forward volatility in foreign exchange. *Journal of Financial Economics* 100(3):496 – 513.

Della Corte, P., T. Ramadorai, and L. Sarno. 2016. Volatility risk premia and exchange rate predictability. *Journal of Financial Economics* 120(1):21 – 40. ISSN 0304-405X.

Engle, R. F., S. Giglio, B. Kelly, H. Lee, and J. Stroebel. 2020. Hedging Climate Change News. *The Review of Financial Studies* 33(3).

Giglio, S., B. Kelly, and J. Stroebel. 2021a. Climate finance. *Annual Review of Financial Economics* 13:15–36. doi: 10.1146/annurev-financial-102620-103311.

Giglio, S., M. Maggiori, K. Rao, J. Stroebel, and A. Weber. 2021b. Climate change and long-run discount rates: Evidence from real estate. *The Review of Financial Studies* 34(8):3527–3571. doi: 10.1093/rfs/hhab032.

Giglio, S., T. Kuchler, J. Stroebel, and X. Zeng. Biodiversity risk. Technical Report w31137 National Bureau of Economic Research 2023 https://www.nber.org/papers/w31137.

Grootendorst, M. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Hassan, T. 2013. Country size, currency unions, and international asset returns. *The Journal of Finance* 68:2269–308.

Hassan, T., T. Mertens, and T. Zhang. 2015. Currency manipulation. *Working Paper*.

Hassan, T. A., T. M. Mertens, and T. Zhang. 2016. Not so disconnected: Exchange rates and the capital stock. *Journal of International Economics* 99:S43–S57.

Hassan, T. A., S. Hollander, L. Van Lent, and A. Tahoun. 2019. Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics* 134(4):2135–2202.

Hassan, T. A., J. Schreger, M. Schwedeler, and A. Tahoun. 2023. Sources and Transmission of Country Risk. *The Review of Economic Studies*. ISSN 0034-6527.

Hijioka, Y., E. Lin, J. J. Pereira, R. T. Corlett, X. Cui, G. Insarov, R. Lasco, E. Lindgren, A. Surjan, E. M. Aizen et al. 2014. Ipcc fifth assessment report: Chapter 24 asia. *Intergovernmental Panel on Climate Change* pages 1327–1370.

Hsu, P.-h., K. Li, and C.-y. Tsou. 2023. The pollution premium. *The Journal of Finance* 78(3):1343–1392.

Javadi, S., A.-A. Masum, M. Aram, and R. P. Rao. 2023. Climate change and corporate cash holdings: Global evidence. *Financial Management* 52(2):253–295.

Kashyap, A., N. Kovrijnykh, and A. Pavlova. 2024. Designing esg benchmarks. *Working paper*.

Lee, S. O., N. C. Mark, J. Nauerz, J. Rawls, and Z. Wei. 2022. Global temperature shocks and real exchange rates. *Journal of Climate Finance* 1:100004.

Leombroni, M., A. Vedolin, G. Venter, and P. Whelan. 2021. Central bank communication and the yield curve. *Journal of Financial Economics* 141(3). ISSN 0304-405X.

Lustig, H., N. Roussanov, and A. Verdelhan. 2011. Common risk factors in currency markets. *Review of Financial Studies* 24:3731–77.

Lustig, H., N. Roussanov, and A. Verdelhan. 2014. Countercyclical currency risk premia. *Journal of Financial Economics* 111(3):527–553.

Mueller, P., A. Stathopoulos, and A. Vedolin. 2017. International correlation risk. *Journal of Financial Economics*.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2021. Sustainable investing in equilibrium. *Journal of Financial Economics* 142(2):550–571. doi: 10.1016/j.jfineco.2020.12.011.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2022. Dissecting green returns. *Journal of Financial Economics* 146(2):403–424. doi: 10.1016/j.jfineco.2022.07.007.

Pavlova, A., and R. Rigobon. 2007. Asset prices and exchange rates. *Review of Financial Studies* 20: 1139–1181.

Pavlova, A., and R. Rigobon. 2010. An asset-pricing view of external adjustment. *Journal of International Economics* 80:144–156.

Pavlova, A., and R. Rigobon. 2013. International macro-finance. *Handbook of Safeguarding Global Financial Stability: Political, Social, Cultural, and Economic Theories and Models* 2:169–176.

Pigato, M. A., and F. E. Stewart. 2020. Climate change in apec: Assessing risks, preparing financial markets, and mobilizing institutional investors. *World Bank Report*.

Reimers, N., and I. Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics 11 2019.

Sandulescu, M., F. Trojani, and A. Vedolin. 2020. Model-free international stochastic discount factors. *Journal of Finance*. forthcoming.

Sautner, Z., L. Van Lent, G. Vilkov, and R. Zhang. 2023. Firm-Level Climate Change Exposure. *Journal of Finance*.

Starks, L. T. 2023. Presidential address: Sustainable finance and esg issues – value versus values. *Journal of Finance* 78:1837–1872. doi: 10.1111/jofi.13255.

Sun, Y., X. Zhang, Y. Ding, D. Chen, D. Qin, and P. Zhai. 2022. Understanding human influence on climate change in china. *National science review* 9(3):nwab113.

Zviadadze, I. 2017. Term structure of consumption risk premia in the cross section of currency returns. *Journal of Finance*.

Chapter 3

The Rise of Green Banks in the U.S.

Sole-authored.[1]

Abstract

I examine how U.S. Green Banks (i.e., public or mission-driven lenders that allocate capital for climate related projects) shape the flow of green finance in the U.S. I assemble project-level microdata for 11 public and quasi-public Green Banks across nine states, and present first evidence that higher borrowing costs and coordination frictions lower the likelihood of private co-financing while raising the Green Bank's dollar contribution per project. A Stackelberg model, with a welfare maximising Green Bank that co-finances with profit-maximising lenders, highlights leverage caps and coordination frictions as key constraints.

Keywords: Green Banks, Climate Finance, Financial Inclusion, Public-Private Partnerships.
JEL Codes: G21, Q55

[1]This chapter is the latest draft of my Job Market Paper. I am deeply grateful to my advisor Max Croce, Riccardo Colacito, Nicola Borri, Roberto Marfè, and the Bocconi Finance faculty for invaluable feedback. I am grateful to Caroline Flammer, Ran Duchin, Sijia Fan, Yongheng Sun, Seung Min Kim, and Yinshan Shang for their thoughtful comments, and to all participants at the NBER Climate Finance Workshop, the FINEST Workshop, and the Bocconi-Bristol Green Finance Workshop for insightful discussions. I gratefully acknowledge financial support from the Baffi Centre at Bocconi University through its *New Frontiers of Green Lending* Ph.D. fellowship, and from the *Fondazione Romeo ed Enrica Invernizzi*. I am also indebted to the many public-sector staff who processed my FOIA requests and provided invaluable data for this project. The latest version and the Internet Appendix are available on my personal webpage.

3.1 Introduction

Green Banks are publicly-backed or quasi-public institutions designed to mobilize private investment for clean energy and climate related projects. Despite their name, Green Banks are not deposit-taking institutions and do not operate as traditional commercial banks. Instead, they function as specialized intermediaries that deploy public funds through loans, credit enhancements, co-investments, and guarantees to address market failures and financing gaps in green sectors. Their core mission is to leverage limited public resources to mobilize larger flows of private capital, especially in contexts where traditional financing has been limited. Over time, Green Banks have also taken on mandates to promote inclusive access to green finance, directing investments toward underserved or disadvantaged communities as part of broader environmental and social policy goals.

This paper examines the rapid institutional expansion of Green Banks in the United States and their role in shaping the flow and structure of green capital. The analysis is motivated by the landmark 2023 allocation of $20 billion by the U.S. Environmental Protection Agency (EPA) through the Greenhouse Gas Reduction Fund (GGRF), representing the largest federal investment to date in these institutions.[2] Despite their growing importance, there remains limited empirical evidence on how Green Banks deploy capital, interact with traditional lenders, and how their activity shapes access to green finance across communities.[3]

The paper addresses three central questions: (1) How have Green Banks evolved institutionally across U.S. states? (2) How is climate finance allocated across regions and sectors through Green Banks? and (3) To what extent do Green Banks expand access to green finance for low- and moderate-income or otherwise underserved communities?

While the policy rationale for Green Banks is that they address financing gaps left by traditional markets, this paper does not attempt to establish whether they increase aggregate private investment in a causal sense. Instead, I focus on the mechanisms through which they could fill such a gap: the allocation of capital across projects, the way risk is shared with private lenders, and the targeting of underserved communities. By documenting these mechanisms, the paper provides the empirical foundation necessary for additional work on the aggregate impact of Green Banks.

To interpret how institutional constraints shape private participation, I develop a model of Green Banks as catalytic intermediaries that co-finance projects with private lenders. The framework

[2] On 11 March 2025 newly appointed EPA Administrator Lee Zeldin issued a press release terminating the $20 billion Greenhouse Gas Reduction Fund (GGRF) awards, alleging "fraud, waste and abuse" (EPA (2025); see also Politico (2025)). Several awardees immediately sued. On 15 April 2025 a preliminary injunction was granted blocking the termination (Post (2025)). The litigation is ongoing and the final outcome remains uncertain.

[3] While working on this draft, I became aware of work by Rizzi et al. (2025), who study the impact of Green Bank presence on venture capital investment in climate-tech startups. Their focus is on startup-level investment dynamics and signaling mechanisms, whereas my paper centers on the institutional evolution and capital deployment strategies of Green Banks, especially in the context of inclusive access to green finance.

highlights how two forms of leverage jointly determine the scale and composition of investment. *Accounting leverage* refers to the use of debt on a Green Bank's own balance sheet, such as tax-exempt bonds or credit lines. *Mobilization leverage* refers to the amount of private capital attracted per dollar of public funding at the project level. These two forms of leverage are linked through the institution's balance sheet constraint. When borrowing capacity is limited or debt is costly, the Green Bank must rely more on private partners to achieve a given project scale. When it can expand its balance sheet, it can finance a larger share internally, which lowers the relative private contribution. The key prediction is that tighter leverage limits or higher borrowing costs reduce the overall mobilization of private capital once both scale and composition effects are taken into account. This mechanism motivates the empirical analysis that follows.

A key contribution of the paper is the construction of a new hand-collected dataset that documents the project-level activity, financial instruments, and sectoral composition of U.S. Green Banks. Because there is no centralized reporting system, I assembled the data from primary sources including audited financial statements, regulatory filings, Freedom of Information Act (FOIA) requests, and internal disclosures obtained directly from the institutions. The resulting dataset links project-level financing, co-investment structures, and location characteristics across major Green Banks, providing the first systematic view of how these institutions deploy capital and engage private partners. Using these data, I show that institutions facing tighter borrowing limits or higher effective leverage costs are less likely to secure private co-financing and tend to fund a larger share of projects with their own capital. Projects that involve more partners, and therefore greater coordination intensity, exhibit systematically lower private participation. Across specifications, both balance-sheet and coordination frictions reduce the equilibrium private share of project financing, consistent with the theoretical mechanism. The evidence shows that balance sheet flexibility and coordination capacity are first-order determinants of how effectively Green Banks mobilize private capital.

These findings highlight the importance of institutional design in shaping the scale and composition of green finance. Green Banks with sustained access to leverage instruments such as bond issuance authority or revolving funds are better positioned to expand lending and attract private participation. Coordination frictions point to the value of institutional support for deal structuring, standardized contracts, and credit-enhancement facilities that can reduce transaction costs in blended finance. Stable, long horizon capitalization also emerges as a critical factor, since episodic or uncertain funding increases effective leverage costs and weakens the catalytic role that motivates Green Bank policy in the first place.

To the best of my knowledge, this is the first paper to build a comprehensive dataset on U.S. Green Banks and connect it to broader questions of institutional evolution and inclusive green finance in the United States.

Beyond these results, the paper contributes a broader empirical foundation for understanding mission-oriented financial institutions in the climate transition. By linking project-level financing, instruments, and sectoral composition across the U.S. Green Bank system, it shows how institutional design shapes the mobilization of private capital and the distributional outcomes of green finance. This evidence provides a basis for future research on the local economic impacts of Green Bank lending, the persistence of crowd-in effects, and the role of institutional form in expanding equitable access to green capital.

Related literature. This paper contributes to three strands of research. First, it relates to the growing literature on sustainable and green finance, which studies how financial markets and institutions respond to the climate transition and how market frictions shape capital allocation (e.g., Flammer (2021), Giglio et al. (2021), Baker et al. (2022), Baldauf et al. (2020), Krueger et al. (2020)). This paper adds to this work by focusing on the institutional architecture that enables the mobilization of private capital through public–private partnerships such as Green Banks.

Second, it connects to the emerging literature on blended finance (e.g., Flammer et al. (2024), Rizzi et al. (2025)), which analyzes how concessional or publicly backed instruments attract private investment into socially valuable but capital-constrained sectors. Flammer et al. (2024) provide the first systematic evidence on blended finance structures globally. My paper complements this work by assembling project-level data on blended finance deals for Green Banks in the United States and by quantifying how leverage constraints and coordination frictions shape private crowd-in within these transactions.

Third, it contributes to the literature on public financial intermediation and development banking, which examines how state-backed or mission-driven financial institutions allocate capital and influence long-horizon investment (e.g., Lazzarini et al. (2015), Mertens et al. (2021), Griffith-Jones and Ocampo (2018)). Green Banks operate within this broader category but often at a subnational level and with more flexible governance arrangements. By documenting how institutional design affects both the scale and distribution of green lending, this paper provides new evidence on the role of public intermediaries in expanding equitable access to climate finance.

Section 3.2 provides institutional context on the emergence and design of U.S. Green Banks. Section 3.3 develops a conceptual framework that formalizes how leverage constraints and coordination frictions shape the mobilization of private capital. Section 3.4 defines the sample, data construction, and project classification. Section 3.4.1 presents summary statistics and descriptive evidence on institutional scale, financial instruments, and sectoral composition. Section 3.4.2 outlines the empirical design and reports the main results linking balance sheet constraints and coordination frictions to private crowd-in. Section 3.5 concludes with policy implications and directions for future research.

3.2 Institutional Background

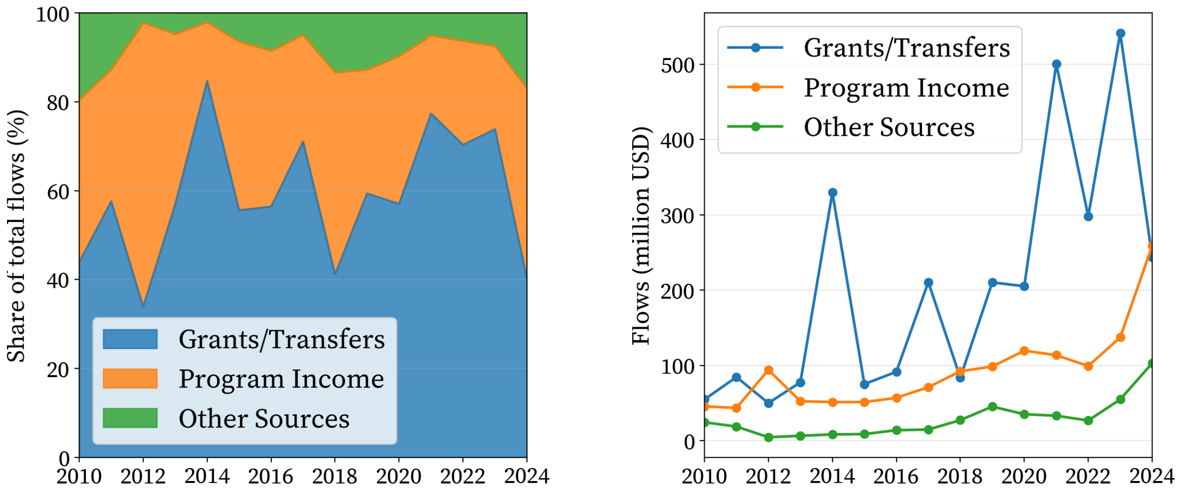
Green Banks emerged in the aftermath of the 2008 financial crisis, when several states sought to scale clean investment without relying exclusively on grants or annual appropriations. The Connecticut Green Bank, established in 2011, is widely recognized as the first U.S. example. Since then, states and localities have created similar institutions, often organized as nonprofit entities or quasi-public agencies. Although they share a common mission, their legal forms, oversight arrangements, and funding models vary widely, reflecting the experimental and decentralized nature of the U.S. approach to green finance.

Legal charters. Four basic forms cover almost every case. (i) *Quasi-public corporations* created by their own state law, such as the Connecticut Green Bank and New York Green Bank. They can issue revenue bonds, offer credit guarantees, and recycle program income. (ii) *State finance authorities* that added a green-bank unit to an older infrastructure or development statute (for example, Rhode Island Infrastructure Bank or California IBank). (iii) *City development authorities* that run green-bank programs under local ordinance (for example, Invest Atlanta or Finance New Orleans); and (iv) *Independent 501(c)(3) nonprofits*, for example Michigan Saves and Inclusive Prosperity Capital, which rely on grants and partner lenders because they cannot tax or issue their own bonds.

Governance and oversight. A state-chartered Green Bank is treated like any other public authority: board seats for state officials, open meetings, yearly budgets approved in a hearing, and audited financials. Nonprofit Green Banks sit in yet another box. They file Form 990 with the IRS, follow state charity law, and must keep donor money in line with the stated mission, but they are not examined for credit risk the way a traditional bank is. Some nonprofits add voluntary safeguards (advisory boards, published loan policies, and third-party impact reviews) to reassure funders.

Funding structure. Because Green Banks are non-depository institutions, they depend on external capitalization and program income rather than deposits. External funding takes two main forms. The first is state transfers, including appropriations, utility system-benefit charges (utility bill surcharges), and bond proceeds. Some of these flows are predictable and recurring (as in Connecticut's use of ratepayer surcharges), while others depend on annual legislative approval and are inherently uncertain. The second category is grants and federal or philanthropic allocations, including proceeds from the Regional Greenhouse Gas Initiative (RGGI), the State Small Business Credit Initiative (SSBCI), Department of Energy programs, and charitable foundations.

Over time, internally generated program income (i.e, loan repayments, interest, and fee-based revenue) has become increasingly important. In Rhode Island, repayments from the Efficient Buildings Fund are recycled into new loans; in New York, NY Green Bank now finances its



(a) Share of revenue **(b) Total revenue (millions USD)**

Figure 3.1: Composition and volume of Green Bank revenues by source.

Notes: The figure reports the composition and total volume of Green Bank revenues, disaggregated by source. Revenues are classified into three categories: (i) transfers and grants (state appropriations, utility surcharges, carbon-auction proceeds, federal allocations, and philanthropic inflows); (ii) program income (loan repayments, interest on promissory notes, sales of renewable energy credits, and fee-based revenues); and (iii) other revenues (investment income and miscellaneous items). Data are aggregated across all Green Banks in the baseline sample and cover fiscal years 2010–2024. The definition of Green Banks follows Section 3.4, which restricts attention to mission-driven lenders that recycle public or philanthropic capital through repayable instruments to crowd in private climate investment.

entire operating budget from earned income. Connecticut and Maryland have also reached the point where internally generated revenue equals or surpasses state transfers. As shown in Figure 3.1, early Green Banks relied heavily on transfers and grants, but program income has matured into a durable revolving base, marking a transition from grant dependence to financial self-sufficiency. This evolution highlights a central friction in Green Bank operations: external capitalization is episodic and limited, constraining leverage expansion, while internal revenues grow only gradually with portfolio maturity.

Conduit issuance. A subset of state authorities (e.g., Ohio Air Quality Development Authority, Illinois Finance Authority) primarily serve as conduit issuers. In this role, the Green Bank (or related authority) issues tax-exempt bonds on behalf of outside borrowers. Although many conduit transactions support private developers, state climate and development finance authorities also use this structure to finance energy and resilience upgrades for municipalities, school districts, transit agencies, and other public or nonprofit borrowers, who can then access tax-exempt debt and standardized documentation without issuing on their own. The bonds are legally and financially the borrower's liability: the Green Bank has no repayment obligation, and bondholders rely on the project's cash flows. From the issuer's perspective, conduit transactions are off balance sheet. They therefore expand the total volume of mobilized green capital without increasing the Green Bank's accounting leverage. Over the past decade, annual issuance has at times exceeded \$3–4 billion in a single state, and outstanding balances for some issuers now reach into the tens of billions. These patterns highlight that, although not part of core on-balance sheet flows, conduit bonds remain a major source of mobilized green capital across several states.

In my database, I track conduit issuance projects alongside on-balance sheet activity. This allows me to capture the full scope of green capital mobilization, while still distinguishing between loans and investments that sit on a Green Bank's balance sheet and those that are purely conduit in nature.

Reporting. State chartered and quasi-public Green Banks must comply with statutory disclosure rules requiring audited financial statements, annual budgets, and regular performance reports. Nonprofit Green Banks primarily disclose through the IRS's Form 990, but many voluntarily publish annual reports, project-level disclosures, and environmental impact metrics. The recent \$20 billion allocation through the U.S. Environmental Protection Agency's Greenhouse Gas Reduction Fund (GGRF) introduces a new layer of standardized federal reporting, especially on greenhouse gas reductions, private capital mobilization, and service to disadvantaged communities.

This fragmented regulatory landscape presents both a challenge and an opportunity for data collection. Because there is no centralized reporting framework, the dataset I construct in this paper is based on original hand collection from a range of primary sources, including institutional reports and mandatory disclosure regulations. This approach allows for detailed project-level analysis across a wide range of institutional models and is described in greater detail in Section 3.4.

3.3 Conceptual Framework

The institutional features described above motivate a conceptual framework for how Green Banks allocate capital and mobilize private finance when operating under balance sheet and coordination frictions. These institutions aim to expand climate investment while facing limited leverage and costly engagement with private partners. Higher leverage enables them to finance more projects but increases borrowing costs and risk exposure. Greater reliance on private co-financing attracts additional capital yet raises the effort required to structure and monitor deals. The model formalizes this trade-off by combining a balance sheet constraint, reflecting the capped and episodic nature of public funding, with a coordination cost that scales with deal complexity. The resulting equilibrium determines how much of each project the Green Bank finances itself and how much private capital is crowded in. The model's predictions about how funding constraints and financing costs affect private participation are then tested empirically using the dataset described in Section 3.4.

3.3.1 Project financing and capital structure

Let L^{GB} denote the amount of capital that the Green Bank invests in a project. The Green Bank selects a co-financing share $\mu \in [0, 1)$, which determines the total project size:

$$L = \frac{L^{GB}}{1 - \mu}. \tag{3.1}$$

The Green Bank finances its investment using tax-exempt debt D and internal funds E (e.g., grants, retained surpluses, or philanthropic contributions), subject to a regulatory leverage constraint:

$$L^{GB} = D + E, \tag{3.2}$$

$$D \leq \lambda L^{GB} \tag{3.3}$$

λ is a statutory cap on the debt share. I assume the cap binds in equilibrium, consistent with convex debt costs and the mission orientation of Green Banks.

Borrowing costs increase with leverage, reflecting political, administrative, or reputational risk:

$$r_D = r_0 + \phi \left(\frac{D}{D + E} \right)^2, \tag{3.4}$$

where r_0 is the base cost of debt and $\phi > 0$ measures cost convexity.

Co-financing introduces coordination frictions that rise with the scale and complexity of private-sector involvement (μ) and project size (L). These are modeled as:

$$F(\mu, L) = \chi \mu^2 L, \tag{3.5}$$

where $\chi > 0$ captures the per-unit marginal cost of managing public–private project delivery.

The Green Bank maximizes a utility function combining mission impact and financial sustainability:

$$\max_{L^{GB}, \mu} \ \theta \log \left(\frac{L^{GB}}{1 - \mu} \right) + (1 - \theta) \log \left[\rho L^{GB} - r_D D - F(\mu, L) \right], \tag{3.6}$$

where $\theta \in (0, 1)$ is the relative weight on mission versus financial performance, $\rho > 0$ is the return per dollar invested, and the remaining terms reflect debt and coordination costs.

The functional forms are chosen for tractability and to capture standard economic trade-offs in a transparent way. The logarithmic specification reflects diminishing returns to both project scale and financial surplus, ensuring an interior solution for the Green Bank's allocation problem. The quadratic debt cost captures the convex increase in borrowing costs as leverage rises, consistent with political and reputational risks that scale nonlinearly with debt exposure. Similarly, the quadratic coordination cost in μ implies that joint project management becomes disproportionately costly as private participation expands, which aligns with observed increases in transaction complexity in large blended-finance deals.

3.3.2 Private bank participation

After observing the Green Bank's proposal (L^{GB}, μ), a private bank (PB) decides whether to provide the complementary capital μL. It earns an expected return per dollar of investment equal to:

$$R^{PB} = \rho - \gamma(\mu, L), \tag{3.7}$$

where γ captures risk-adjusted coordination costs and is increasing in the scale of private participation. I model coordination costs as rising in total deal scale, reflecting fixed transaction costs and the complexity of multi-party negotiations. Coordination costs for the Private Bank are modelled as:

$$\gamma(\mu, L) = \psi \mu L \tag{3.8}$$

with $\psi > 0$ denoting the sensitivity of private bank costs to the scope of their involvement. The private bank invests only if the risk-adjusted return exceeds its required return r^{PB}:

$$\rho - \gamma(\mu, L) \geq r^{PB} \tag{3.9}$$

This yields a participation constraint:

$$\mu \cdot \frac{L^{GB}}{1-\mu} \leq \frac{\rho - r^{PB}}{\psi} \tag{3.10}$$

The participation constraint states that the private bank will only co-invest if the expected return from the project, net of coordination costs, exceeds its required rate of return. The right-hand side represents the maximum total scale of joint financing that the private sector is willing to undertake, given its opportunity cost of capital (r^{PB}) and the sensitivity of coordination costs (ψ). A higher required return or stronger sensitivity to deal complexity tightens the constraint, limiting the feasible size of co-financed projects. Conversely, projects with higher expected returns (ρ) relax the constraint and allow greater private participation for a given Green Bank commitment.

Definition 3.3.1 (*Equilibrium*)**.** The Stackelberg equilibrium is a pair (L^{GB*}, μ^*) such that the Green Bank maximizes its utility function (3.6), subject to the following constraints and (3.1):

$$\begin{aligned} &L^{GB} = D + E, \\ &D \leq \lambda L^{GB}, \\ &\mu \cdot \frac{L^{GB}}{1-\mu} \leq \frac{\rho - r^{PB}}{\psi} \quad \text{(Private bank participation constraint)} \end{aligned}$$

The Green Bank acts as a Stackelberg leader: it anticipates the private bank's response to any co-financing proposal (L^{GB}, μ) and chooses its strategy accordingly. The equilibrium characterizes the optimal allocation of public capital and co-investment share, ensuring participation by private banks while balancing the Green Bank's mission and financial objectives.

Proposition 3.3.1 (*Stackelberg Equilibrium in Green Co-Financing*)**.** Let the Green Bank choose (L^{GB}, μ) to maximize (3.6) under (3.2)–(3.10). Then the unique interior equilibrium is:

$$\mu^* \text{ solves } \mu^2(2\theta - 1)\chi + \mu\theta A - A = 0, \quad 0 < \mu^* < 1,$$

and

$$L^{GB*} = K\,\frac{1-\mu^*}{\mu^*}, \qquad D^* = \lambda\, L^{GB*}, \qquad L^* = \frac{L^{GB*}}{1-\mu^*},$$

where

$$\bar{r}_D = r_0 + \phi\lambda^2, \quad A = \rho - \lambda\bar{r}_D, \quad K = \frac{\rho - r^{PB}}{\psi} > 0,$$

with $A > 0$, $\rho > r^{PB}$, $\theta > \frac{1}{2}$, and

$$\chi(2\theta - 1) > (1-\theta)A. \tag{$*$}$$

The equilibrium exists under a set of mild and economically intuitive conditions. The requirement $\rho > r^{PB}$ ensures that projects are sufficiently profitable for private lenders to consider participation. The condition $A > 0$, where $A = \rho - \lambda\bar{r}_D$, guarantees that the expected project return exceeds the Green Bank's effective cost of debt, making public financing sustainable. The parameter restriction $\theta > \frac{1}{2}$ implies that the Green Bank places greater weight on mission impact than on short-term financial returns, which is consistent with its mandate as a catalytic public intermediary. Finally, the inequality $\chi(2\theta - 1) > (1-\theta)A$ ensures an interior equilibrium where both public and private contributions are positive. Intuitively, coordination costs cannot be so low that the Green Bank crowds in all private capital, nor so high that it becomes optimal to finance projects entirely on its own.

Lemma 3.3.1 (*Comparative statics of μ^* in χ and λ*)**.** Let μ^* solve the GB's optimization problem, with $\theta \in (\frac{1}{2}, 1)$, $\chi > 0$, $A > 0$, $r_0 > 0$, $\phi > 0$, and $\lambda > 0$. Then:

$$\frac{\partial \mu^*}{\partial \chi} < 0, \qquad \frac{d\mu^*}{d\lambda} < 0.$$

Proposition 3.3.1 pins down a unique interior co-financing share in this model. As shown in Lemma 3.3.1, μ^* decreases in both the marginal coordination cost of coordination, χ, and the GB's leverage ratio cap, λ. The comparative statics are intuitive: a higher coordination-friction parameter makes each extra unit of crowd-in more costly, so the GB chooses a lower crowd-in share ($\partial \mu^* / \partial \chi < 0$). Because K is fixed, a lower μ mechanically requires a larger project size L^* and, therefore, more GB capital, L^{GB}. Similarly, raising GB's leverage, λ increases the GB's debt cost r_D and shrinks the surplus A; with less surplus to "share," the optimal crowd-in share falls ($\partial \mu^* / \partial \lambda < 0$), again implying a larger project L^* and heavier GB balance-sheet usage. Policywise, lowering co-financing frictions ($\downarrow \chi$) or debt cost convexity ($\downarrow r_0$, $\downarrow \phi$) would mobilize the same private dollars with less GB capital tied up per project, freeing GB capacity to back more deals.

Testable implication. If debt becomes more expensive or constrained, the Green Bank optimally lowers the private crowd-in share (μ) while keeping the absolute amount of private capital roughly constant, since the private bank's participation constraint binds. Empirically, when the Green Bank faces higher leverage costs or tighter borrowing limits, we should observe a decline in the co-financing share μ and an increase in the Green Bank's own contribution per project, while private dollars per deal remain relatively stable within short windows where the private bank hurdle K does not change. These patterns imply that balance-sheet tightening shifts the composition of financing toward the Green Bank without reducing total project size. The elasticity test follows from the model identities $L^{GB} = K(1 - \mu)\mu$ and $L = K/\mu$, which link observed deal composition to the underlying crowd-in parameter. Estimating how the ratio of public to private financing responds to variation in leverage costs or borrowing authority provides a direct test of the model's central mechanism: balance-sheet flexibility determines how efficiently Green Banks can mobilize private capital for a given level of total investment.

3.4 Data

Definition of Green Bank. To construct a consistent list of Green Banks, I apply a working definition grounded in both legal structure and operational function. Specifically, I define Green Banks as mission-driven financial institutions, typically quasi-public agencies or nonprofit entities, that deploy public or philanthropic capital through repayable financial instruments such as loans, credit enhancements, and co-investments to mobilize private investment in clean energy and climate resilience. This definition excludes climate-focused institutions that engage solely

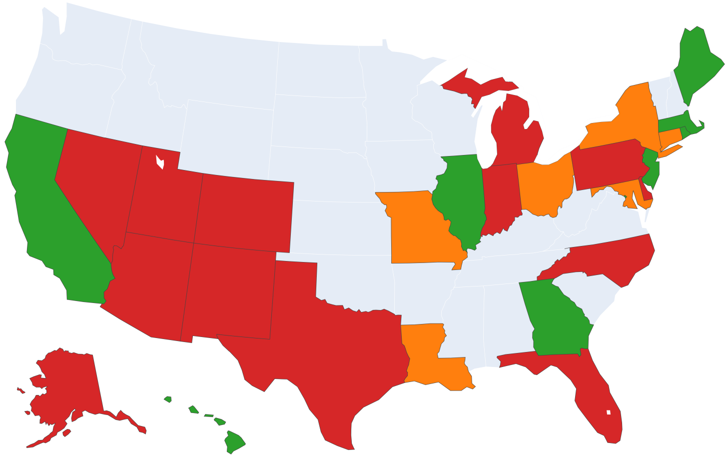


Figure 3.2: U.S. states that host at least one Green Bank, 2011–2024.

Notes: Shaded jurisdictions meet the Green Bank definition from Section 3.4 ("mission-driven lenders that recycle public or philanthropic capital through repayable instruments to crowd in private climate investment"). Highlighting is based on the list in Table 3.1 and reflects the institutional landscape as of July 2025. Both public/quasi-public agencies and nonprofit green banks are included; states hosting more than one qualifying institution (e.g., California, Massachusetts, Ohio) are plotted in orange. Puerto Rico, which also hosts a green bank, lies outside the map frame and is omitted.

in education, advocacy, or grantmaking, as these organizations do not function as financial intermediaries or revolve capital.

The primary sources for identifying candidate institutions are the Coalition for Green Capital, a national nonprofit that advocates for and incubates Green Bank models, and the Green Bank 50 coalition, an informal network of roughly 50 green lending institutions that formed in 2024 to share data and coordinate access to new federal programs such as the EPA's GGRF.

Institutions are classified as Green Banks if they meet at least three of the following criteria: (i) their core mission includes mobilizing climate-related private capital; (ii) they use market-based financial tools; (iii) they are structured to allow capital recycling; and (iv) they operate under a public mandate or nonprofit governance model. This filtering strategy ensures the dataset captures institutions that actively shape green capital flows through direct financial activity. Table 3.1 summarizes the list of Green Banks in my sample.

Data collection. A central contribution of this paper is the construction of a novel dataset on Green Bank activity in the United States. Because no centralized reporting system exists, the dataset was built from the ground up through extensive hand collection from multiple primary sources, including audited financial statements, Form 990 filings, EMMA-posted official statements, program dashboards, and project-level registries. For the roughly half of Green Banks organized as public or quasi-public authorities, I submitted state open-records (FOIA-equivalent) requests to obtain detailed internal data. Private nonprofit Green Banks are not subject to FOIA, so I relied on direct outreach to institutional staff. Many of these non-profit entities were in early development stages and could only share partial information, which required merging heterogeneous files and reconciling inconsistent accounting definitions across states and years. Where detailed project records were unavailable, I hand-recorded balance sheet and program aggregates from the most recent public disclosures. This intensive collection effort produced a harmonized dataset that captures the financial activity of Green Banks across a wide range of institutional models and reporting practices.

Table 3.1: U.S. Green Banks: Legal Form and Year Founded

Name	State	Year	Type	Sub-type
Connecticut Green Bank	Connecticut	2011	Public Agency	—
New York Green Bank	New York	2013	Public Agency	—
Rhode Island Infrastructure Bank	Rhode Island	1989	Public Agency	—
California Infrastructure and Economic Development Bank (IBank)	California	1994	Public Agency	—
Hawaii Green Infrastructure Authority (HGIA)	Hawaii	2013	Public Agency	—
DC Green Bank	District of Columbia	2018	Public Agency	—
California Alternative Energy and Advanced Transportation Financing Authority (CAEATFA)	California	1980	Public Agency	—
California Pollution Control Financing Authority (CPCFA)	California	1972	Public Agency	—
Atlanta Development Authority *d/b/a* Invest Atlanta	Georgia	1979	Public Agency	—
Illinois Finance Authority / Climate Bank	Illinois	2004	Public Agency	—
Finance New Orleans	Louisiana	1978	Public Agency	—
Efficiency Maine Trust	Maine	2010	Public Agency	—
Maryland Clean Energy Center	Maryland	2008	Public Agency	—
Massachusetts Clean Energy Center	Massachusetts	2008	Public Agency	—
Massachusetts Community Climate Bank	Massachusetts	2023	Public Agency	—
Environmental Improvement and Energy Resources Authority (EIERA)	Missouri	1972	Public Agency	—
New Jersey Economic Development Authority	New Jersey	1974	Public Agency	—
Ohio Air Quality Development Authority	Ohio	1970	Public Agency	—
Michigan Saves	Michigan	2009	Non-profit	Public Charity
Montgomery County Green Bank	Maryland	2016	Non-profit	Public Charity
Energize Delaware	Delaware	2010	Non-profit	Supporting Org.
Solar and Energy Loan Fund of Florida (SELF)	Florida	2010	Non-profit	Public Charity
Nevada Clean Energy Fund	Nevada	2017	Non-profit	Public Charity
Colorado Clean Energy Fund	Colorado	2018	Non-profit	Public Charity
Louisiana Clean Energy Fund	Louisiana	2023	Non-profit	Public Charity
Indiana Energy Independence Fund	Indiana	2023	Non-profit	Public Charity
Missouri Green Banc	Missouri	2022	Non-profit	Private Foundation
New Mexico Climate Investment Center	New Mexico	2023	Non-profit	Public Charity
New York City Energy Efficiency Corporation (NYCEEC)	New York	2010	Non-profit	Public Charity
North Carolina Clean Energy Fund	North Carolina	2020	Non-profit	Public Charity
Columbus Region Green Fund	Ohio	2021	Non-profit	Public Charity
GO Green Energy Fund (Growth Opps)	Ohio	2020	Non-profit	Public Charity
Philadelphia Green Capital Corp.	Pennsylvania	2021	Non-profit	Public Charity
Puerto Rico Green Energy Trust	Puerto Rico	2019	Non-profit	Public Charity
Clean Energy Fund of Texas	Texas	2021	Non-profit	Public Charity
SustainEnergyFinance	Utah	2023	Non-profit	Public Charity
Inclusive Prosperity Capital	National (CT)	2018	Non-profit	Public Charity
Spruce Root	Alaska	2012	Non-profit	Public Charity
Groundswell Capital	Arizona	2022	Non-profit	Supporting Org.

Notes: “Public Agency” covers both state agencies and quasi-public authorities created by statute or local ordinance. “Non-profit” entries follow IRS determinations; “Supporting Org.” is a 509(a)(3) structure. Years are the enabling statute date or, for non-profits, the IRS incorporation year. Source: author compilation from statutes, Form 990 filings, and institutional websites.

Financial information. All balance-sheet and income figures come from publicly-available primary documents; no secondary databases are used. I rely on five source types that together cover every Green Bank in the sample:

(i) *State Comprehensive Annual Financial Reports (CAFRs)* and *component-unit audits*. Quasi-public and state-agency green banks appear in the enterprise-fund or component-unit sections of their state CAFR; the PDFs include a full Statement of Net Position, a Statement of Revenues and Expenses, and all debt footnotes.

(ii) *Standalone audited financial reports.* About one-third of the quasi-public banks (e.g. Connecticut Green Bank, DC Green Bank) publish their own GAAP-based audit separate from the CAFR. These reports provide detail on program loans, loan-loss reserves, and restricted net position that is not in the CAFR.

(iii) *IRS Form 990 filings*. Every non-profit bank files Form 990; the balance sheet (Part X) and revenue statement (Part VIII) supply cash, grants receivable, secured debt, and unrestricted net assets. XML versions are pulled from the ProPublica Non-profit Explorer webpage; PDFs are used when XML is missing.

(iv) *MSRB EMMA disclosures*. Whenever a Green Bank issues tax-exempt or "green" bonds, the Official Statement and continuing-disclosure filings on EMMA list coupon rates, amortisation schedules, and outstanding principal. These numbers are cross-checked against the liabilities shown in (i)–(iii).

(v) *Federal-grant and program reports*. For institutions that receive EPA Greenhouse Gas Reduction Fund (GGRF) or other public funds, I confirm draw-downs and unspent balances from the first-quarter 2024 financial reports, posted on USASpending.gov.

If a figure conflicts across documents, I follow a simple hierarchy: standalone audit → CAFR note → Form 990 → EMMA. Missing items (mostly for start-up non-profits that have not yet filed a 990) are requested directly from staff or, where necessary, via state FOIA. A log of every PDF and XML file, together with the code used to parse them, is archived in the Internet Appendix for reproducibility.

In this paper, I define a Green Bank's leverage as the ratio of its debt obligations (e.g., tax-exempt bonds, credit lines, or other borrowing) to total capital (debt plus equity-type funds such as grants, retained earnings, and philanthropic contributions). For banks that operate solely on grants or appropriations, leverage is zero.[4] This definition allows consistent comparison across

[4]Several Green Banks in the sample (particularly nonprofit start-ups) operate entirely on grants, appropriations, or philanthropic equity and carry no debt. For these institutions, leverage is mechanically zero in all years. In the baseline regressions, I code their leverage ratio as zero, which both preserves the full sample and reflects their institutional reality. Results are robust to (i) excluding non-leveraged banks, and (ii) interacting leverage with a nonprofit indicator. Both approaches yield qualitatively similar coefficients, suggesting that the estimated effect of leverage constraints is not driven solely by banks with borrowing authority.

institutions with different legal forms and funding models.

Project-level information. I collect project-level information for 11 Green Banks across 9 U.S. states between 2011 and 2024.[5] Each observation contains project characteristics such as funding amount, financial instrument (loan, guarantee, credit enhancement), co-financing details, target sector (e.g., residential solar, commercial efficiency), and geographic location (ZIP or census tract level where available). I use three primary sources:

(a) *Annual reports*: every Green Bank that publishes an Annual Impact Report, ACFR, or "Transaction Profile" series is scraped for deal-level tables.

(b) *Board packets and sunshine-law portals*: meeting minutes and project registers released under state open-records statutes (e.g., NY Open Meetings Law, Connecticut FOI) provide the most granular line-item data.

(c) *Direct requests and FOIA/FOIL queries*: for banks without public deal logs, redacted spreadsheets are obtained via negotiated data-sharing agreements or formal FOIA requests to the state energy office [ongoing].

Project classification. To align the empirics with the theoretical model, I focus on transactions where Green Banks commit capital that is at risk and repayable. Pure grants and subsidies are excluded from the baseline sample, since they do not revolve, carry no repayment obligation, and bypass leverage or co-financing frictions. Hybrid structures that combine subsidies with repayable loans are coded only on the repayable component. Contingent risk-sharing instruments such as funded loan loss reserves are retained, as they represent capital deployed and exposed to potential loss.

All transactions are then classified into four mutually exclusive categories: (i) *on-balance lending*, such as direct loans, leases, participations, or mezzanine financing from Green Bank funds; (ii) *credit enhancements*, such as cash reserves, or collateral support that are fully funded and at risk; (iii) *conduit issuance*, when the Green Bank issues tax-exempt bonds on behalf of a private borrower, but credit risk remains with the borrower and the transaction is off-balance sheet; and (iv) *facilitation-only*, which are projects fully financed by private lenders where the Green Bank provides only convening or technical assistance.

[5] The sample covers the following institutions: Connecticut Green Bank; New York Green Bank; Rhode Island Infrastructure Bank; California Infrastructure and Economic Development Bank (IBank); California Pollution Control Financing Authority (CPCFA); Illinois Finance Authority (Climate Bank program); Maryland Clean Energy Center; Massachusetts Clean Energy Center; Massachusetts Community Climate Bank; Environmental Improvement and Energy Resources Authority (EIERA, Missouri); and New Jersey Economic Development Authority. **Ongoing FOIA and data-sharing requests cover the Hawaii Green Infrastructure Authority (HGIA), DC Green Bank, the California Alternative Energy and Advanced Transportation Financing Authority (CAEATFA), the Atlanta Development Authority (Invest Atlanta), Finance New Orleans, Efficiency Maine Trust, and the Ohio Air Quality Development Authority.**

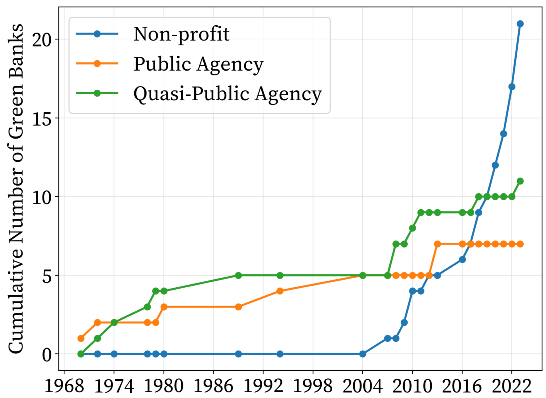
In the baseline regressions, I include categories (i) and (ii), which involve Green Bank dollars at risk and are directly subject to balance-sheet constraints. Conduit issuance and facilitation-only projects are excluded from the baseline but analyzed separately in robustness checks.

Other variables considered. I supplement these project-level data with county-level socioeconomic and climate attitude variables. Median income is drawn from the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) dataset at the county level. Measures of population and GDP are obtained from the U.S. Bureau of Economic Analysis dataset on personal income and economic activity by county. Finally, climate concern is measured using the Yale Program on Climate Change Communication's (YPCCC) county-level survey estimates (Howe et al. (2015)), specifically the share of adults who are somewhat or very worried about global warming (with a national average of 63% in 2024). These sources allow me to examine whether co-financing is systematically related to both project-level financial structure and broader county-level socioeconomic and climate context.

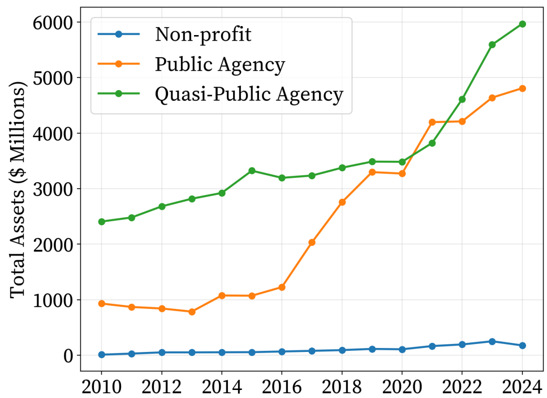
Lastly, I classify projects by sector using the Climate Bonds Initiative (CBI) taxonomy. This taxonomy is an internationally recognized classification system developed by CBI to identify investments consistent with climate mitigation and resilience goals. It provides standardized sector categories (e.g., solar energy, electric transport) and is widely used in green bond certification. Using this taxonomy ensures comparability of project types across institutions and time, and aligns my classification with the broader sustainable finance literature.

3.4.1 Summary Statistics

Green Banks have expanded rapidly in both number and balance sheet size over the past decade, though their growth has taken place through distinct institutional forms. Figure 3.3 illustrates this evolution. Panel (a) shows that while early initiatives were primarily established as public or quasi-public agencies, the number of nonprofit Green Banks has accelerated sharply since 2015, reflecting a broader diversification of institutional models. Panel (b) documents a parallel increase in total assets, particularly among quasi-public and public agencies, whose combined balance sheets now exceed $10 billion. Nonprofit institutions remain smaller in scale but are growing quickly, often focusing on local or community-based lending. These patterns highlight both the dynamism and heterogeneity of the U.S. Green Bank ecosystem, which highlights the need to analyze project-level outcomes and leverage behavior at the institutional level.



(a) Cumulative number of Green Banks



(b) Total assets (millions USD)

Figure 3.3: Growth and balance sheet expansion of Green Banks in the U.S.

Notes: This figure summarizes the institutional development and financial growth of Green Banks in the United States. Panel (a) reports the cumulative number of Green Banks founded in the United States, disaggregated by legal structure. Institutions are classified according to their organizational form (e.g., nonprofit entities, public or quasi-public agencies) as reported in Table 3.1. Year of establishment is taken from official Green Bank filings, press releases, or institutional reports, and the sample reflects the universe of qualifying institutions active between 2010 and 2024. The definition of Green Banks follows Section 3.4, which restricts attention to mission-driven lenders that recycle public or philanthropic capital through repayable instruments to crowd in private climate investment. Panel (b) shows the evolution of total assets (in millions of USD) held by active Green Banks over the same period, by legal structure. Asset data are compiled from annual reports, audited financial statements, or equivalent disclosures.

Table 3.2: Summary statistics of selected green banks and related institutions.

Green Bank / Institution	Proj.	FY Min	FY Max	Cap. (M)	Proj. Amt. (M)	Avg. (M)	Med. (M)	Share Cof. (%)	Pub/Priv. Cof. (%)	# Ctys	Top Sector
Atlanta Development Authority (Invest Atlanta)	16	2011	2025	3.87	58.74	3.67	1.73	68.8	58.9	2	Buildings
California Infrastructure & Economic Dev. Bank	97	2000	2024	634.58	1000.52	10.31	5.69	81.4	22.9	34	Water
Connecticut Green Bank	10,321(*)	2013	2025	157.93	742.41	0.07	0.02	100.0	95.1	8	Buildings
DC Green Bank	35	2020	2024	53.13	186.29	5.32	1.80	100.0	18.4	1	Energy
Environmental Improvement & Energy Resources Authority	22	2010	2025	0.00	1022.91	46.50	26.04	100.0	100.0	5	Water
Illinois Finance Authority / Climate Bank	27	2024	2026	28.90	40.80	1.51	0.45	100.0	14.9	16	Energy
Maryland Clean Energy Center	551	2013	2025	0.32	193.34	0.35	0.01	100.0	94.8	19	Buildings
Massachusetts Clean Energy Center	56	2018	2024	9.16	12.11	0.22	0.10	100.0	5.0	6	Energy
Massachusetts Community Climate Bank	49	2025	2026	3.40	3.40	0.07	0.07	100.0	0.0	13	Buildings
NY Green Bank	86	2016	2025	1358.94	1538.44	17.89	10.00	4.7	3.1	20	Energy
New Jersey Economic Dev. Authority	48	2010	2024	35.83	35.83	0.75	0.21	—	0.0	0	Land Use
Ohio Air Quality Dev. Authority	212	2010	2023	36.54	4422.27	20.86	1.17	94.8	94.8	64	Industry
Rhode Island Infrastructure Bank	229	2015	2024	1026.04	1142.65	4.99	0.94	100.0	15.7	5	Water

Notes: The table reports summary statistics for major U.S. Green Banks. "Committed capital" is the total amount directly deployed on balance sheet by the institution (USD millions). "Project amount" includes both committed capital and mobilized co-financing (USD millions). "Avg." and "Med." refer to the average and median project size. "Share co-financed" is the percentage of projects with any co-financing reported. "Public/Private co-financed" is the percentage of projects that involve both public and private sources of capital. "# Ctys" indicates the number of distinct counties in which projects have been located. "Top sector" identifies the sector with the largest share of projects. (*) The very large project count primarily reflects its *Smart-E* program, under which thousands of small retail loans for home energy upgrades (and related measures) are originated via partner lenders and recorded as individual projects. Source: author's database, compiled from institutional reports and bond disclosures.

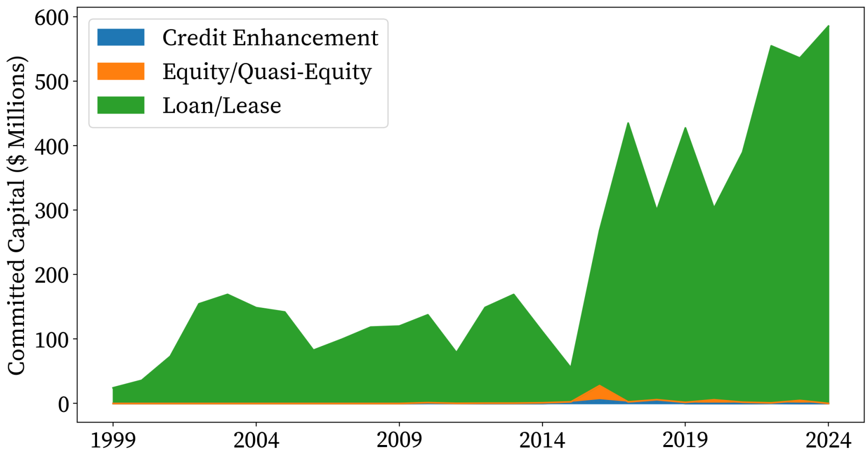


Figure 3.4: Committed Capital by Instrument Type, U.S. Green Banks.

Notes: This figure reports the aggregate dollar volume of capital commitments made by U.S. Green Banks between fiscal years 2010 and 2024, disaggregated by instrument type. "Committed capital" refers to the face value of financing approved for deployment through loans, leases, credit enhancements, and other repayable instruments. Conduit bond issuances are excluded to focus on balance-sheet or directly intermediated activity. Instrument types are mutually exclusive and follow the harmonized taxonomy in Section 3.4. Fiscal years correspond to each institution's own reporting calendar. All amounts are shown in nominal dollars.

Table 3.2 summarizes the scope and activities of major U.S. Green Banks and related state authorities. The institutions vary widely in scale: while Connecticut Green Bank has financed over 10,000 projects with relatively small average ticket sizes, NY Green Bank and Rhode Island Infrastructure Bank have mobilized over a billion dollars each through larger, fewer transactions. Patterns of co-financing also differ: some institutions, such as the Environmental Improvement and Energy Resources Authority and Maryland Clean Energy Center, report nearly universal private participation, while others, like NY Green Bank, focus on balance-sheet lending with limited co-investment. The sectoral distribution likewise reflects institutional design, with state infrastructure banks (California, Rhode Island, Ohio) concentrating on water and industrial facilities, while clean energy focused banks (Connecticut, DC, Massachusetts) primarily target buildings and solar energy.

Figure 3.4 shows that Green Bank activity is dominated by loans and leases throughout the sample, with credit enhancements and equity/quasi-equity appearing only intermittently and at much smaller scales. Aggregate committed volumes are modest in the early years, then rise sharply from the mid-2010s onward, with notable surges around the late 2010s and again in the early 2020s. The composition remains stable even as totals grow: lending expands while equity and guarantees remain secondary tools. Interpreted jointly, the pattern is consistent with Green Banks operating primarily as balance-sheet lenders that recycle capital through repayable instruments rather than as equity investors, while deploying targeted credit enhancements episodically. Because conduit bond programs are excluded, the figure focuses on directly intermediated activity and therefore understates total mobilization channels that run through

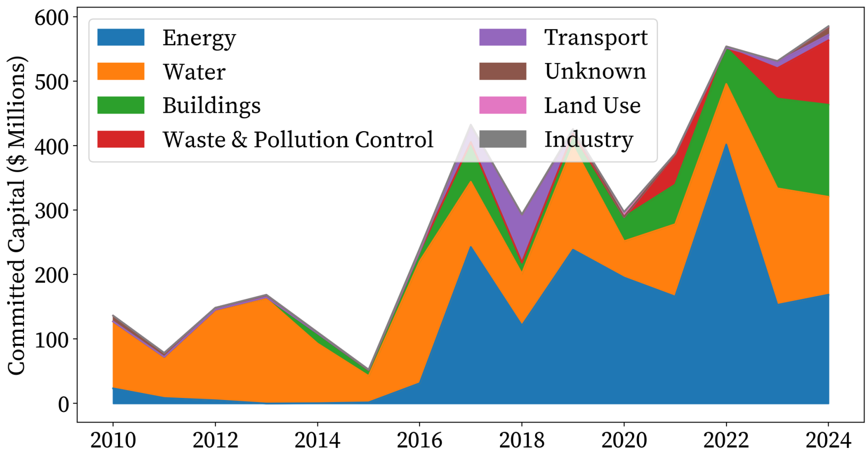


Figure 3.5: Committed capital (Loan/Lease) by sector, 2010–2024.

Notes: The figure shows annual committed capital aggregated across Green Banks in the sample, restricted to instruments classified as `Loan/Lease`. The stacked areas split totals by sector of the project (based on the CBI Taxonomy). "Conduit Debt" is excluded. Fiscal years follow each institution's reporting calendar. Amounts are nominal USD. If a sector is missing for a transaction, it is grouped under "Unknown."

conduit structures.

Figure 3.5 highlights how the composition of Green Bank lending shifted over time, even though loans and leases dominate overall activity. In the early years, most loan volume is concentrated in water-related projects, typically involving infrastructure upgrades and treatment systems. Beginning around 2016, there is a clear pivot toward energy, reflecting the growing use of standardized clean energy financing models adopted across institutions. In the 2020s, there is also a visible expansion of building sector lending, driven largely by retrofit and efficiency improvements rather than new construction. Activity in waste, land use, transport, and industrial segments appears only intermittently and at modest scales, often linked to isolated remediation or facility upgrade efforts. The sharp increases in total volume after the late 2010s stem primarily from intensified lending within the energy and building sectors, consistent with the broadening mandates and programmatic expansion of state and nonprofit lenders rather than a shift into fundamentally new areas.

Figure 3.6 shows the private share of total project capital by fiscal year. For each year, the share equals the sum of project totals minus Green Bank committed amounts, divided by the sum of project totals, excluding "Conduit Debt" to focus on balance-sheet or directly intermediated financing. Higher values indicate more non–Green Bank capital relative to total costs. GBs show a clear upward trend in the private share of total project capital over the last 14 years. Interpreted literally, these dynamics suggest a growing ability of Green Bank activity to crowd in non-Green Bank resources over time, alongside sensitivity to the mix and timing of large transactions. Because "Conduit Debt" is excluded, the figure focuses on balance-sheet or directly intermediated activity and omits mobilization channeled through conduit bond programs.

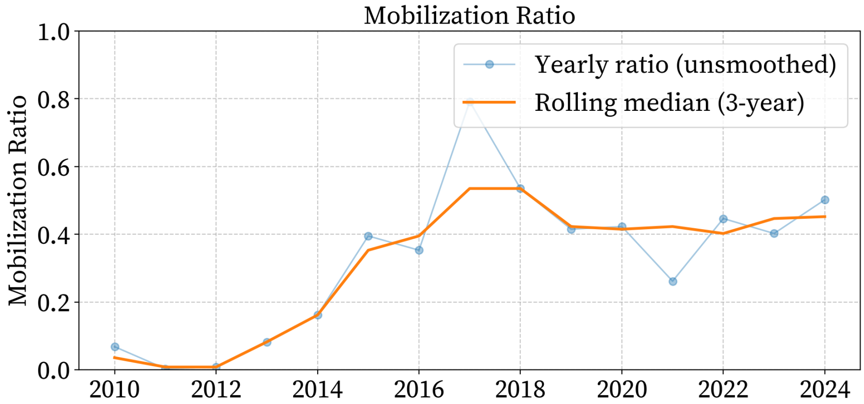


Figure 3.6: Mobilization leverage.

Notes: The series plots the aggregate mobilization leverage for the GB in my sample, i.e., the private share of total project capital by fiscal year. For each year, this share is computed as the sum of project totals minus the Green Bank's committed amounts, divided by the sum of project totals. The numerator captures non-Green Bank (private and other non-GB) capital mobilized. "Conduit Debt" projects are excluded to focus on balance-sheet or directly intermediated activity. When the total project cost is missing, it is treated as the GB's committed capital for the construction of the series. Fiscal years follow each institution's reporting calendar. Amounts are nominal, and ratios are transaction-sum weighted (not simple averages).

Finally, when a project's total cost is missing I set it equal to the Green Bank's committed amount; this assigns a zero private share to those observations and makes the aggregate series a conservative (lower-bound) estimate of private mobilization.

3.4.2 Empirical results

Extensive and Intensive Margins of Co-Financing

The model described in Section 3.3 predicts that the Green Bank's optimal private crowd-in share, $\mu^{\star}$, declines when (i) balance-sheet leverage becomes more costly and (ii) coordination frictions with private lenders increase. Empirically, I map these two theoretical frictions to observable proxies that vary across institutions and projects.

Leverage costs. In the model, tighter or more expensive leverage (higher λ or higher r_D) constrains the Green Bank's capacity to scale lending. In the data, I proxy this constraint using the Green Bank *accounting leverage ratio*, defined as total debt divided by total assets. A higher ratio reflects greater balance-sheet usage and higher marginal borrowing costs, capturing how close an institution is to its effective leverage ceiling. Because capitalization events are episodic, this ratio provides a natural summary measure of available balance-sheet slack at time t.

Coordination frictions. The model's coordination cost, χ, captures the effort required to originate, structure, and monitor projects jointly with private lenders. I proxy this friction using the number of distinct non-GB participants associated with each project, such as commercial banks, investors, or community development financial institutions (CDFIs). Projects involving multiple private partners tend to require more negotiation and oversight, reflecting higher coordination intensity. This proxy captures variation in deal complexity and alignment costs across projects and Green Banks.

Empirically, I observe both (i) whether a project secures any private co-financing (participation, i.e., the extensive margin) and (ii) the intensity of private participation relative to public dollars when it occurs (the intensive margin). I therefore estimate a two-part empirical design that maps the model's comparative statics to observable outcomes.

Part I: Extensive margin (participation). Define the indicator

$$\text{Cofinanced}_{ibt} \in \{0, 1\},$$

which equals one if project i at green bank b in year t involves at least one private lender. I estimate the probability of crossing this participation hurdle as a function of leverage costs and coordination frictions:

$$\Pr(\text{Cofinanced}_{ibt} = 1) = F\Big(\alpha + \beta_1\, \text{LeverageCost}_{bt} + \beta_2\, \text{Coordination}_{ibt} + \beta_3' X_{ibt} + \gamma_b + \delta_t\Big), \tag{3.11}$$

where $F(\cdot)$ is a logit or probit CDF; X_{ibt} includes project controls (CBI sector dummies and instrument type); and γ_b, δ_t are green bank and year fixed effects. Standard errors are clustered at the bank level. Consistent with the Stackelberg logic, the model implies $\beta_1 < 0$ and $\beta_2 < 0$.

As a robustness check allowing for high-dimensional fixed effects (e.g., GB$\times$state or county), I also estimate a linear-probability model:

$$\text{Cofinanced}_{ibt} = \alpha + \beta_1\, \text{LeverageCost}_{bt} + \beta_2\, \text{Coordination}_{ibt} + \beta_3' X_{ibt} + \gamma_b + \delta_t + \varepsilon_{ibt}, \tag{3.12}$$

with heteroskedasticity-robust standard errors clustered at the bank level. The signs of $\widehat{\beta}_1, \widehat{\beta}_2$ should mirror the marginal effects from (3.11).

Part II: Intensive margin (private share). Conditional on participation, the model implies that tighter or more costly leverage and greater coordination frictions induce the Green Bank to finance a larger share of each project, while the private partner's contribution is pinned by its participation constraint. Consequently, the *private share of total financing* (μ_{ibt}) should decline

Table 3.3: Intensive Margin: Private Share vs. Leverage and Coordination Frictions

	(1) Pooled OLS	**(2) OLS + FE**	**(3) Within-GB**
Leverage Ratio	-0.171	-0.136	-0.526**
	(0.182)	(0.331)	(0.188)
Coordination Cost	-0.257***	-0.303***	-0.431***
	(0.059)	(0.078)	(0.029)
Year FE	No	Yes	Yes
Instrument FE	No	Yes	Yes
Green Bank FE	No	No	Yes (demeaned)
Controls	Yes	Yes	Yes
Observations	498	498	498
R^2	0.043	0.162	0.101

Notes: The dependent variable $\mu_{ibt} = \frac{\text{Private}_{ibt}}{\text{Private}_{ibt}+\text{GB}_{ibt}}$ measures the private share of total project financing (intensive margin). "Leverage Ratio" proxies for Green Bank balance-sheet tightness, and "Coordination Cost" captures the complexity and alignment effort associated with private partners. Models are estimated by OLS with standard errors clustered by Green Bank (in parentheses). Column (1) reports pooled OLS; Column (2) includes year and instrument fixed effects; Column (3) is a within-bank specification that demeans continuous variables at the Green Bank level. Negative coefficients indicate that higher leverage or coordination frictions are associated with lower private participation, consistent with the model's prediction that tighter constraints reduce the optimal private share $\mu^\star$. Stars denote significance at the 10% (*), 5% (**), and 1% (***) levels.

as leverage costs or coordination frictions increase. I therefore measure the intensive margin as

$$\mu_{ibt} \;=\; \frac{\text{Private}_{ibt}}{\text{Private}_{ibt} + \text{GB}_{ibt}},$$

and estimate

$$\mu_{ibt} = \alpha' + \theta_1 \,\text{LeverageCost}_{bt} + \theta_2 \,\text{Coordination}_{ibt} + \theta_3' X_{ibt} + \gamma_b' + \delta_t' + \varepsilon_{ibt}'. \tag{3.13}$$

Interpretation. Higher values of LeverageCost_{bt} correspond to tighter or more expensive effective leverage in the sense of the model, and higher $\text{Coordination}_{ibt}$ captures greater alignment costs with private partners. Under this mapping, the model predicts $\theta_1 < 0$ and $\theta_2 < 0$ in (3.13): when either friction increases, the Green Bank optimally absorbs a larger share of financing, reducing the equilibrium private share $\mu^\star$. For completeness, when the outcome is the Green Bank's *dollar contribution* per co-financed deal, the same comparative statics imply positive coefficients on these variables.

I report (i) pooled OLS, (ii) OLS with year and instrument fixed effects, and (iii) a within-bank specification that demeans continuous regressors and the dependent variable at the Green Bank level, absorbing bank-specific averages.

Signs and statistical significance. Across specifications, both the *Leverage Ratio* and *Coordination Cost* coefficients are negative, and the latter is highly statistically significant throughout. This pattern is consistent with the model's comparative statics: tighter balance-sheet constraints and higher coordination frictions lead the Green Bank to finance a larger share of each project, reducing the equilibrium private share of total financing, μ_{ibt}. The magnitude of the leverage coefficient increases and becomes statistically significant in the within-bank specification (Model 3), indicating that variation in leverage over time within a Green Bank is associated with lower private participation. The robustness of the coordination-cost effect across all models further suggests that deal complexity and alignment with private partners are first-order determinants of crowd-in intensity.

Economic magnitude. All ratios (e.g., leverage) are expressed in decimal form; thus, coefficients correspond to effects per unit change in the ratio. Interpreting Model (3), a 0.1 (10 percentage point) increase in a Green Bank's leverage ratio is associated with a 0.053 point decline in the private share of total project financing. Likewise, a one-unit increase in the coordination-cost proxy reduces the private share by roughly 0.43. Given that the average private share μ_{ibt} is below one-half, these magnitudes are economically meaningful: moderate tightening of leverage or increases in coordination complexity substantially reduce private crowd-in at the project level. The within-bank evidence thus reinforces the model's prediction that tighter balance-sheet conditions or higher coordination costs depress the optimal private share $\mu^{\star}$ even within the same institution over time.

Link to the participation (extensive) margin. The two-part logic maps closely to the theory. Higher leverage costs and coordination frictions reduce the probability that any private lender participates ($\beta_1, \beta_2 < 0$). Conditional on participation, tighter constraints also lower the private share of financing, as Green Banks fund a larger fraction of each project themselves ($\theta_1, \theta_2 < 0$). Empirically, the intensive-margin result in Table 3.3 aligns with this mechanism.

Overall, the evidence indicates that tighter GB financing conditions are associated with both a lower likelihood of private participation (extensive margin) and a lower intensity of private capital per public dollar (intensive margin). These patterns are consistent with the model's comparative statics and suggest that policies which relax leverage constraints or reduce coordination frictions (standardized contracts, arranger support, or blended-finance facilities) can increase private crowd-in.

Identification. A concern in estimating equations (3.11)–(3.13) is that the Green Bank's leverage cost, LeverageCost_{bt}, is endogenous to private mobilization outcomes. Projects that attract little private capital may leave the Green Bank carrying a larger share of the financing, mechanically raising leverage. Conversely, favorable project environments could reduce both leverage and the private share, confounding the relationship.

To recover the causal effect of financing constraints on crowd-in, I plan to exploit plausibly exogenous variation in balance-sheet tightness arising from funding shocks and institutional rules. These include (i) the timing of bond issuances and scheduled debt maturities, which affect available borrowing capacity independently of contemporaneous project demand; (ii) budget reallocations or capital injections from state legislatures, which shift the Green Bank's leverage potential but are predetermined with respect to specific projects; and (iii) regulatory or statutory changes in debt authority that alter allowable leverage caps λ across time or institutions. Together, these shocks generate within-bank variation in effective leverage costs that is orthogonal to project-level fundamentals.

In parallel, I will strengthen the identification of coordination frictions by using measures of project complexity and partner composition that are determined prior to financing outcomes. For example, the presence of specialized intermediaries (arrangers, credit enhancers) can serve as exogenous proxies for alignment costs that are not mechanically linked to private funding volumes. These strategies will allow me to isolate the causal impact of financing constraints and coordination frictions on private crowd-in, directly testing the mechanisms highlighted in Proposition 3.3.1.

3.5 Conclusions

This paper provides the first systematic, data-driven analysis of how U.S. Green Banks mobilize private capital under balance-sheet and coordination frictions. Combining a conceptual framework with newly hand-collected project-level data, I show that leverage constraints and coordination costs play a central role in shaping both the likelihood and intensity of private participation in green finance deals.

The theoretical framework formalizes Green Banks as catalytic intermediaries that maximize a joint objective combining mission impact and financial sustainability under regulatory and organizational constraints. The model yields sharp comparative statics: tighter leverage limits or higher borrowing costs lead to lower optimal private crowd-in shares, while greater coordination frictions reduce both participation and scale. The empirical results strongly support these predictions. Together, these findings reveal that balance-sheet flexibility and deal coordination capacity are first-order determinants of how effectively Green Banks crowd in private capital.

The evidence has three main implications. First, Green Banks with sustained access to leverage instruments, such as bond issuance authority or revolving funds, are better positioned to expand lending and mobilize private participation. Second, the coordination friction estimates highlight the value of institutional support for deal structuring: standardized contracts, technical assistance funds, and credit-enhancement facilities can materially lower the cost of blended finance arrangements. Third, the results highlight the need for stable, long-horizon capitalization. Episodic funding cycles increase effective leverage costs, weakening the catalytic function that motivates Green Bank policy in the first place.

Methodologically, this paper contributes a new empirical foundation for studying mission-oriented financial institutions in the climate transition. The dataset assembled here, which links project-level financing, instruments, and sectoral composition across the major U.S. Green Banks, creates opportunities for future work on several open questions: the local economic impacts of Green Bank lending, the persistence of mobilization effects over time, and the role of institutional design in shaping equitable access to green capital.

Looking forward, the identification strategies outlined in the previous section, will allow sharper causal tests of how exogenous funding shocks or regulatory changes influence private crowd-in. Extending this framework to incorporate community-level outcomes and environmental performance metrics would further clarify how Green Banks can enhance both financial and social returns in the green transition.

In sum, Green Banks have emerged as pivotal intermediaries in the architecture of U.S. climate finance. By quantifying how their institutional constraints shape the mobilization of private capital, this paper highlights both the promise and the policy design challenges of building a scalable, inclusive, and financially sustainable green finance system.

Bibliography

Baker, M., D. Bergstresser, G. Serafeim, and J. Wurgler. 2022. Financing the response to climate change: The pricing and ownership of us green bonds. *Brookings Papers on Economic Activity*.

Baldauf, M., L. Garlappi, and C. Yannelis. 2020. Does climate change affect real estate prices? only if you believe in it. *Review of Financial Studies* 33(3):1256–1295.

EPA. Administrator zeldin terminates $20 billion in ggrf awards 2025 https://www.epa.gov/newsreleases/administrator-zeldin-terminates-biden-harris-20b-gold-bar-grants.

Flammer, C. 2021. Corporate green bonds. *Journal of Financial Economics* 142(2):499–516. doi: 10.1016/j.jfineco.2021.05.008.

Flammer, C., T. Giroux, and G. M. Heal. Blended finance. Technical Report 973/2024 European Corporate Governance Institute (ECGI) Brussels 2024. 62 pp.; posted April 2 2024; last revised September 14 2025; Available at SSRN: https://ssrn.com/abstract=4770779.

Giglio, S., M. Maggiori, J. Stroebel, and A. Weber. 2021. Climate change and long-run discount rates: Evidence from real estate. *Review of Financial Studies* 34(8):3527–3571.

Griffith-Jones, S., and J. A. Ocampo, editors. *The Future of National Development Banks*. Oxford University Press Oxford 2018. ISBN 9780198827948.

Howe, P. D., M. Mildenberger, J. R. Marlon, and A. Leiserowitz. 2015. Geographic variation in opinions on climate change at state and local scales in the usa. *Nature Climate Change* 5(6):596–603. doi: 10.1038/nclimate2583.

Krueger, P., Z. Sautner, and L. T. Starks. 2020. The importance of climate risks for institutional investors. *Review of Financial Studies* 33(3):1067–1111. doi: 10.1093/rfs/hhz137.

Lazzarini, S. G., A. Musacchio, R. Bandeira-de Mello, and R. Marcon. 2015. What do state-owned development banks do? evidence from BNDES, 2002–09. *World Development* 66:237–253. doi: 10.1016/j.worlddev.2014.08.016.

Mertens, D., M. Thiemann, and P. Volberding. *The Reinvention of Development Banking in the European Union: Industrial Policy in the Single Market and the Emergence of a Field*. Oxford University Press Oxford 2021. ISBN 9780198859703.

Politico. Epa's zeldin terminates $20b in biden climate grants 2025 https://www.politico.com/news/2025/03/11/zeldin-terminates-greenhouse-gas-reduction-fund-grants-00225481.

Post, W. Judge blocks epa from taking back $14 billion in climate grants 2025 https://www.washingtonpost.com/climate-environment/2025/04/16/trump-epa-climate-grants-lawsuit/.

Rizzi, C., S. Xu, and P. Yoo. Government-funded green banks: Catalysts for the green transition. European Finance Association (EFA) 2025 Conference Paper 2025 https://papers.ssrn.com/sol3/papers.cfm?abstract$_i d = 5174254$.