Daily Fiscal Stress Signals and Month-End Monitoring: Evidence from Puerto Rico Restructured GO Bonds

Jorge A. Arroyo
Independent Researcher
arroyo.jorgeantonio@gmail.com

February 2026

Abstract

This paper constructs a transparent daily fiscal stress indicator from thinly traded Puerto Rico restructured general obligation (GO) bonds using transaction-level yields from EMMA, benchmarked to matched U.S. Treasury yields from FRED. The index applies an explicit thin-market protocol with capped carry-forward and coverage diagnostics. Because Puerto Rico's monthly Economic Activity Index (EAI/IAE-PR) is not generally available at month-end in real time, the paper evaluates the indicator's usefulness for real-time month-end monitoring (m|m) and tests whether simple distributed-lag month-end mappings provide incremental information for EAI levels and month-over-month changes relative to AR(1) persistence benchmarks. Distributed-lag mappings are estimated on a pre-specified post-reset burn-in overlap—which is necessarily short in the available sample ($N = 9$ monthly observations once lag structure is imposed)—and, with coefficients held fixed, assessed on validation and out-of-sample windows. Results indicate economically interpretable, regime-sensitive monitoring content concentrated in the early post-reset price-discovery episode, but no stable improvements in out-of-sample point prediction accuracy relative to the benchmark. All transformations, evaluation windows, and inputs are reproducible via a pinned AS-OF date, configuration snapshots, and input hashes.

1 Introduction

Small, data-lagged economies often face limited availability of timely macroeconomic indicators. When key aggregates are released with delay or revision, real-time monitoring at month-end (m|m) must rely on alternative information sets. Financial market prices provide one such set: they are observed at high frequency and embed forward-looking assessments of economic conditions and policy constraints (Andreou et al., 2013; Faust et al., 2013). However, for enclosed or administratively constrained economies—including islands, territories, and single-hub jurisdictions—local credit markets may be thin, and extracting a usable daily signal requires explicit treatment of infrequent trading and stale prices (Green et al., 2007; Craig et al., 2018).

This paper examines whether transaction-implied credit spreads from post-restructuring Puerto Rico general obligation (GO) bonds provide a useful real-time monitoring signal and whether that signal exhibits incremental month-end (m|m) mapping content for Puerto Rico's monthly Economic Activity Index (EAI) (Department of Economic Development and Commerce, 2026). Puerto Rico provides a post-restructuring environment in which newly issued or restructured GO bonds begin trading with refreshed credit pricing following a debt settlement effective March 15, 2022 (Medioli et al., 2022; U.S. Government Accountability Office, 2025). The setting is representative of a broader "credit reset" episode in which restructuring, oversight, or capital-structure revision reopens price discovery and concentrates information into market spreads (Chirinko et al., 2018). The empirical question is whether such a spread-based stress signal contains real-time month-end information for monthly activity measures, and whether any such content is stable or regime-dependent (Fuertes and Kalotychou, 2007).

The main object of interest is a daily fiscal stress index constructed from transaction-level yields for three restructured Puerto Rico GO maturities (2035, 2037, 2046) relative to maturity-matched U.S. Treasury yields. For bond j on day t, the fiscal spread is defined as

$$s_{j,t} = y_{j,t}^{\mathrm{PRGO}} - y_{m(j),t}^{\mathrm{UST}}, \tag{1}$$

where $y_{j,t}^{\mathrm{PRGO}}$ is the transaction-implied yield and $y_{m(j),t}^{\mathrm{UST}}$ is the mapped Treasury benchmark. A daily composite is formed as the median across available spreads and then standardized using a rolling procedure computed without look-ahead. Construction is intentionally transparent and auditable: it relies on EMMA transaction data obtained via manual export and on Treasury yields from FRED (Chirinko et al., 2018; Chalmers et al., 2019), with an explicit thin-trading protocol (stale/no-trade classification, capped carry-forward) and accompanying coverage diagnostics (Getmansky et al., 2004; Qian, 2011).

The month-end evaluation asks whether monthly aggregates of the daily fiscal stress index improve real-time month-end (m|m) inference about (i) the level of EAI and (ii) month-over-month changes in EAI, relative to a minimal persistence benchmark. The baseline predictive specification is a distributed-lag regression of the form

$$y_m = \alpha + \beta_0 RT_m + \beta_1 RT_{m-1} + \varepsilon_m, \tag{2}$$

where y_m denotes either EAI_m or ΔEAI_m, and RT_m is the monthly aggregation of the daily stress index (Andreou et al., 2013; Gomez-Zamudio and Ibarra, 2017). Month-end predictions for month m use only information available as of the end of month m (m|m), and evaluation is conducted under pre-specified burn-in, validation, and out-of-sample windows. Concretely, the month-m regressor RT_m is computed using only daily observations dated within month m (including protocol-governed carry-forward fills) and is never revised using information from month $m+1$ or later.

The paper's empirical motivation is consistent with evidence that sovereign and sub-sovereign default risk is closely linked to real activity and that spreads can reflect default anticipation (Manasse et al., 2003; Chari et al., 2017). In Puerto Rico, increases in GO spreads and default probabilities have been associated with subsequent declines in economic activity, measured using the same EAI series examined here (Chari et al., 2017). At the same time, the municipal bond market is characterized by decentralized over-the-counter trading, sparse quoting, and infrequent transactions, which can produce stale pricing and induce serial correlation in observed series (Green et al., 2007; Chalmers et al., 2019; Lo and MacKinlay, 1989). These features motivate the paper's emphasis on explicit thin-trading rules and diagnostics rather than implicit interpolation.

The contribution is methodological and evaluative. First, the paper constructs a daily, transaction-based fiscal stress index from locally relevant post-restructuring securities, with an explicit thin-trading protocol and accompanying coverage diagnostics (Getmansky et al., 2004; Qian, 2011). Second, it evaluates real-time month-end mapping performance using strict window discipline and benchmark comparisons designed for short samples and potentially nested models, where out-of-sample performance can be sensitive to estimation noise (Clark and West, 2007; Hubrich and West, 2010). Third, the analysis frames results in terms of regime-sensitive information content rather than assuming constant gains: predictive relationships are assessed across windows and interpreted as potentially stronger during early post-restructuring price discovery and weaker in later stabilization periods (Chirinko et al., 2018).

Results indicate economically interpretable within-sample associations between fiscal stress and EAI, with evidence that lagged spreads often contain more information than purely contemporaneous co-movement in the early overlap period and that this pattern is consistent across the three CUSIPs. Out-of-sample performance, however, is mixed: the distributed-lag model does not consistently dominate the AR(1) benchmark across evaluation windows. Windowed performance metrics and rolling diagnostics suggest that information content varies over time, consistent with episodic relevance of post-reset spreads rather than stable improvements in point month-end predictions. In additional checks, augmenting the fiscal signal with local equity components does not materially improve out-of-sample performance relative to the fiscal-only specification.

A practical motivation for this approach is the informational environment of Puerto Rico and similar jurisdictions. Official reporting delays can be substantial, limiting the usefulness of fiscal accounts for real-time assessment (U.S. Government Accountability Office, 2025). The EAI itself is used by local institutions as a summary indicator of economic conditions and is not always available contemporaneously (Departamento de Hacienda de Puerto Rico, 2023; Department of Economic Development and Commerce, 2026). In such settings, transparent market-implied indicators may complement delayed macro releases, provided that their construction is explicit and their performance is evaluated conservatively.

The remainder of the paper proceeds as follows. Section 2 describes the EMMA and FRED data inputs and details the construction of the fiscal spread index, including the thin-trading protocol and diagnostics. Section 3 presents the month-end evaluation design, benchmark models, information-set timing, window discipline, and evaluation metrics. Section 4 reports results, including burn-in diagnostics, out-of-sample month-end mapping performance for EAI levels and changes, and evidence on regime sensitivity. Section 5 discusses interpretation and portability to other post-reset, data-lagged economies. Section 6 concludes.

2 Data and Signal Construction

This section describes the data inputs and the construction of the daily fiscal stress index used in the month-end monitoring/mapping exercise. The goal is to produce a transparent, transaction-based spread series that is locally relevant, auditable, and implementable under thin trading. Signal construction is performed without using information from the target variable (EAI), and all transformations are defined prior to estimation and evaluation.

Notation (core objects).

- Bond index $j \in \{2035, 2037, 2046\}$ and day t.
- $\mathcal{T}_{j,t}$: set of trades for bond j on day t.
- $y^{\text{PRGO}}_{j,t}$: daily transaction-implied yield for bond j (VWAP across $\mathcal{T}_{j,t}$).
- $y^{\text{UST}}_{m(j),t}$: Treasury yield for maturity bucket $m(j)$ on day t.
- $s_{j,t} = y^{\text{PRGO}}_{j,t} - y^{\text{UST}}_{m(j),t}$: daily fiscal spread.
- S_t: daily composite spread (median across valid $s_{j,t}$).
- $\tilde{S}_t$: EWMA-smoothed composite spread (span 10) computed from the carry-forward-completed S_t.
- RT_t: standardized daily fiscal stress signal, $RT_t = (\tilde{S}_t - \mu_{t-1})/\sigma_{t-1}$, where $(\mu_{t-1}, \sigma_{t-1})$ are rolling moments computed over a 252-trading-day window using information through $t-1$.
- Monthly aggregates are denoted by the subscript m (e.g., RT_m), defined in Section 3.

2.1 Data sources and samples

2.1.1 Puerto Rico GO bonds (EMMA; manual export)

The local credit inputs are transaction-level observations for three restructured Puerto Rico general obligation (GO) bonds issued under the Series 2022A restructuring settlement. The sample begins on March 15, 2022, the effective date of the GO settlement (Medioli et al., 2022; U.S. Government Accountability Office, 2025). Trades are obtained from the MSRB's Electronic Municipal Market Access (EMMA) system via manual export.[1]

The three CUSIPs and stated maturities are:

- 2035: 74514L3L9
- 2037: 74514L3M7
- 2046: 74514L3P0

For each trade, the inputs used in construction are: (i) trade date, (ii) trade par amount, and (iii) trade yield (or equivalently price converted to yield under the EMMA convention) (Chalmers et al., 2019; Chirinko et al., 2018). The index is based on transaction-implied

[1] Manual export is used to preserve an auditable acquisition step. The import routine is deterministic and documented. No automated scraping is used.

yields rather than quoted yields, reflecting the decentralized over-the-counter structure of the municipal market (Green et al., 2007; Craig et al., 2018). Thin trading and stale pricing are treated explicitly below.

2.1.2 Treasury benchmarks (FRED)

The external benchmark curve is the U.S. Treasury constant maturity series from FRED. Daily yields are obtained for DGS7, DGS10, and DGS20. These series are used as deep-market references for term structure movements and as a normalization device for local spread extraction, following the maturity-matched Treasury benchmarking approach used in the Puerto Rico GO literature (Chari et al., 2017). The Treasury series are aligned to the fiscal-trade date grid; when a Treasury observation is missing on a given trade date, the most recent prior Treasury yield is carried forward.

The primary maturity mapping is:

$$\begin{aligned} 2035 &\to \text{DGS7}, \\ 2037 &\to \text{DGS10}, \\ 2046 &\to \text{DGS20}. \end{aligned}$$

Section 4 reports robustness checks to alternative mapping choices.

2.1.3 Target series (monthly EAI)

The month-end evaluation targets are the Puerto Rico Economic Activity Index (EAI) and its month-over-month change. The EAI is a widely used summary indicator of economic conditions in Puerto Rico and is used by local institutions to contextualize fiscal outcomes (Departamento de Hacienda de Puerto Rico, 2023). The paper treats the EAI as a delayed macro target and evaluates whether a daily market-implied stress measure provides useful within-month information for month-end monitoring/mapping (m|m); the index definition and construction are taken from the current official methodology (Department of Economic Development and Commerce, 2026).

2.1.4 Data summary

Table 1 summarizes the samples, frequencies, and basic coverage properties. The summary includes the number of daily observations, the frequency of zero-trade days, and mean coverage diagnostics used in the thin-trading protocol.

Table 1: Data summary.

Series	Freq.	Sample period	N	% zero	Med. run	Max run
PR GO 2035	Daily	2022-03 to 2026-02	1024	13.5	1d	4d
PR GO 2037	Daily	2022-03 to 2026-02	1024	14.6	1d	15d
PR GO 2046	Daily	2022-03 to 2026-02	1024	13.7	1d	12d
UST DGS7/10/20	Daily	2010-01 to 2026-02	4205	—	—	—
EAI, ΔEAI	Monthly	1999-07 to 2025-11	317	—	—	—

Notes: "Run" denotes consecutive no-trade days within each GO series.

2.2 Fiscal spread construction

2.2.1 Daily VWAP yields

For each CUSIP j and day t, a single daily yield is computed using a volume-weighted average of transaction-level yields:

$$y_{j,t}^{\mathrm{PRGO}} = \frac{\sum_{k\in\mathcal{T}_{j,t}} q_k\, y_k}{\sum_{k\in\mathcal{T}_{j,t}} q_k}, \tag{3}$$

where $\mathcal{T}_{j,t}$ denotes the set of trades for bond j on day t, y_k is the trade yield, and q_k is the trade par amount. If trade sizes are unavailable for a given export, the daily yield is computed as the simple average of trade yields for (j,t). When no trades occur for (j,t), $y_{j,t}^{\mathrm{PRGO}}$ is missing and handled via the thin-trading protocol in Section 2.3.

2.2.2 Spread definition and daily composite

Daily fiscal spreads are computed by subtracting the mapped Treasury yield:

$$s_{j,t} = y_{j,t}^{\mathrm{PRGO}} - y_{m(j),t}^{\mathrm{UST}}, \tag{4}$$

where $m(j)$ denotes the mapping from bond j to the Treasury maturity bucket. A daily composite spread is formed as the median of the available bond-level spreads:

$$S_t = \mathrm{median}\left(s_{2035,t},\ s_{2037,t},\ s_{2046,t}\right), \tag{5}$$

computed over the subset of bonds deemed valid (non-stale) on day t as defined by the thin-trading protocol. Using the median reduces sensitivity to bond-specific microstructure noise and occasional idiosyncratic prints.

2.2.3 Smoothing and standardization

The composite spread S_t is smoothed using an exponentially weighted moving average (EWMA) to reduce day-to-day microstructure noise in thin markets. Let $\tilde{S}_t$ denote the

Figure 1: Daily fiscal spread signal (raw composite and standardized stress measure).

EWMA-smoothed series computed from the carry-forward-completed composite (Section 2.3). The implementation uses EWMA span 10 (i.e., $\alpha = 2/(10+1)$) with the standard recursion

$$\tilde{S}_t = \alpha S_t + (1-\alpha)\tilde{S}_{t-1}. \tag{6}$$

The primary daily stress signal is then standardized using a rolling z-score computed without leakage:

$$RT_t = \frac{\tilde{S}_t - \mu_{t-1}}{\sigma_{t-1}}, \tag{7}$$

where $(\mu_{t-1}, \sigma_{t-1})$ are rolling moments computed over a fixed trailing window of 252 trading days using only information available through day $t-1$ (rolling_t1). The standardized signal is defined once the full 252-day window is available; prior values are left missing. Higher values of RT_t correspond to higher market-implied fiscal stress.

Figure 1 plots the raw composite spread and its standardized counterpart.

2.3 Thin-trading protocol and diagnostics

Because municipal bonds trade infrequently, observed transaction prices and yields can be stale, and naive interpolation can induce serial correlation and measurement error (Lo and MacKinlay, 1989; Getmansky et al., 2004; Choi et al., 2021). The protocol below constructs a continuous daily composite while explicitly flagging no-trade, stale, and carry-forward events. This mirrors the emphasis in the municipal microstructure literature on sparse trading and delayed incorporation of information (Craig et al., 2018; Chalmers et al., 2019).

2.3.1 Zero-trade rules

The protocol distinguishes between "no-trade" (no bonds trade) and "stale" (only one bond trades in a multi-bond configuration). In the multi-bond setting used in the main analysis (three bonds configured), the daily composite is constructed as follows:

1. If at least two bonds trade on day t, set S_t equal to the median of spreads across traded bonds.

2. If exactly one bond trades on day t, classify the day as *stale* and set S_t to missing prior to carry-forward.

3. If zero bonds trade on day t, classify the day as *no-trade* and set S_t to missing prior to carry-forward.

In a single-bond configuration (used only in robustness checks), the stale classification is not defined and the protocol reduces to a no-trade rule for that bond.

2.3.2 Carry-forward cap and implementation

Missing composites due to stale or no-trade days are filled by carrying forward the most recent valid S_t, subject to a cap on consecutive carry days:

$$\texttt{MAX\_FISCAL\_CARRY\_DAYS} = 20. \tag{8}$$

If the cap is exceeded, the observation remains missing until a valid multi-bond trading day occurs. This cap limits the extent to which a prolonged absence of trades can mechanically smooth the series.

2.3.3 Missing-data handling summary (Treasury vs fiscal carry-forward)

Two distinct carry-forward rules are used. For the Treasury benchmark, missing observations on a fiscal trade date are filled by carrying forward the most recent prior Treasury yield, ensuring that spread construction in (4) is defined on the EMMA trade-date grid. For the fiscal composite, missing S_t values arising from stale or no-trade days (Section 2.3.1) are filled by carrying forward the most recent valid multi-bond composite, subject to the cap in (8). The two rules operate on different objects (benchmark yields vs local composite spreads) and are applied prior to EWMA smoothing and rolling standardization.

2.3.4 Diagnostics

Each day is accompanied by diagnostic flags and summary statistics that quantify data availability and the extent of carry-forward. The main diagnostics include:

- `coverage_ratio`: fraction of configured bonds that traded on day t
- `is_no_trade`: indicator for zero-trade days
- `is_stale`: indicator for one-trade days in multi-bond mode
- `is_carry_event`: indicator for carry-forward fills
- `single_cusip_mode`: indicator for single-bond construction mode

Figure 2: Coverage ratio (fraction of configured bonds trading each day).

Table 2: Thin-trading diagnostics for the fiscal signal (daily; EMMA trades).

Diagnostic	Value	Notes
Zero-trade days (%)	9.47	No configured bonds trade
Stale days (%)	3.52	Only one bond trades (multi-bond mode)
Median carry duration (days)	1.0	Consecutive carry-forward length (applied)
Max consecutive carry (days)	4	Capped at MAX_FISCAL_CARRY_DAYS
Mean coverage ratio	0.861	Average fraction trading

Figure 2 plots the daily coverage ratio. Table 2 reports summary thin-trading diagnostics, including the frequency of zero-trade and stale days and the distribution of carry durations.

3 Nowcast Design and Evaluation Framework

This section describes the month-end (m|m) targets, predictive specifications, window discipline, and evaluation metrics. The design follows a standard out-of-sample predictive evaluation workflow in which models are estimated on an initial burn-in window and evaluated on validation and out-of-sample windows using only information available at the month-end decision point (m|m) (West, 2006; Fuertes and Kalotychou, 2007). The objective is not to maximize in-sample fit, but to assess whether the fiscal stress signal provides incremental real-time month-end *mapping/monitoring* content relative to minimal univariate benchmarks, recognizing that any gains may be episodic in a post-reset price-discovery setting.

3.1 Targets

The analysis considers two monthly targets evaluated at the month-end origin:

1. the level of the Puerto Rico Economic Activity Index, EAI_m, and
2. the month-over-month change, $\Delta EAI_m \equiv EAI_m - EAI_{m-1}$.

Mapping the level assesses whether the signal tracks broad persistence and medium-run movements in activity, while mapping ΔEAI_m focuses on short-horizon fluctuations. Figures 3 and 4 plot the target series.

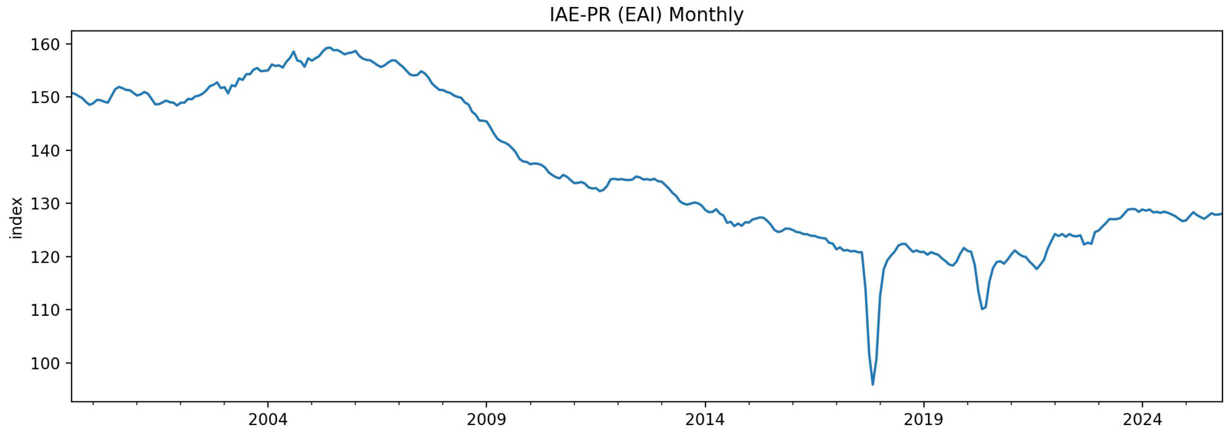

Figure 3: Puerto Rico Economic Activity Index (EAI), monthly level.

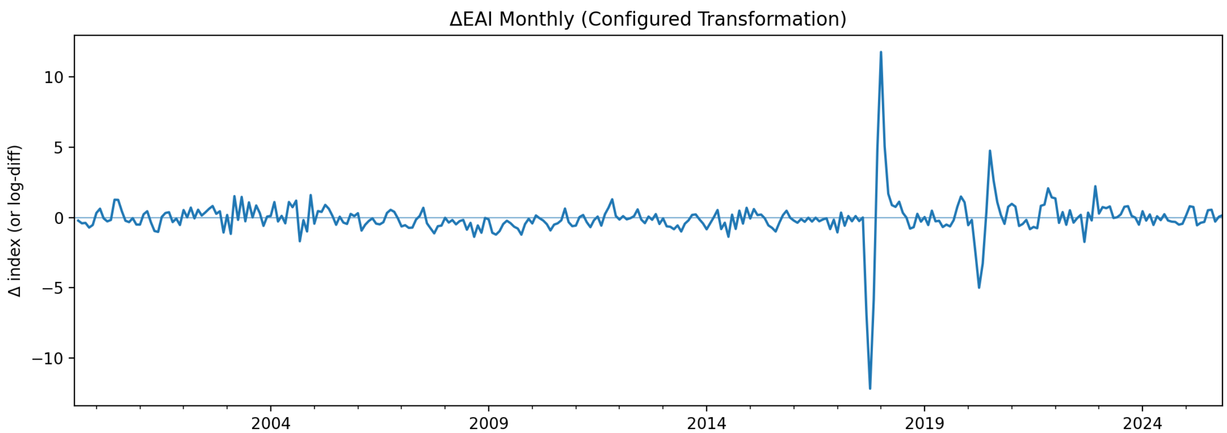

Figure 4: Puerto Rico Economic Activity Index (EAI), month-over-month change.

3.2 Predictive specifications

3.2.1 Daily-to-monthly aggregation

The fiscal stress signal is constructed daily and mapped to the monthly target frequency via flat aggregation. Specifically, the monthly regressor is the within-month mean of the daily standardized stress measure:

$$RT_m = \frac{1}{|D_m|} \sum_{t \in D_m} RT_t, \tag{9}$$

where D_m is the set of days in month m for which the daily stress signal is defined.[2] Flat aggregation is a common baseline in mixed-frequency nowcasting settings, although it may discard intra-month timing information relative to MIDAS-type specifications (Gomez-

[2]In practice, RT_t is defined on the daily date index of the constructed fiscal signal, which includes carry-forward fills governed by the thin-trading protocol.

Zamudio and Ibarra, 2017; Ghysels et al., 2004, 2006; Andreou et al., 2013).

3.2.2 Distributed-lag regression

The primary predictive specification is a distributed-lag (DL) regression with contemporaneous and one-lagged monthly stress:

$$y_m = \alpha + \beta_0 RT_m + \beta_1 RT_{m-1} + \varepsilon_m, \tag{10}$$

where y_m denotes either EAI_m or ΔEAI_m. The DL form allows the stress indicator to affect the target within the same month and with a one-month delay. Opposite-signed coefficients ($\beta_0 > 0$, $\beta_1 < 0$) are interpreted mechanically as a tilt toward a change-like or "innovation" component of the stress index rather than the stress level alone, consistent with the possibility that aggregation and thin-trading frictions can induce persistence that the lag term partially nets out. Coefficient estimates and inference are interpreted cautiously given the short post-restructuring sample (U.S. Government Accountability Office, 2025) and potential serial correlation inherited from thin trading and smoothing (Lo and MacKinlay, 1989; Getmansky et al., 2004; Choi et al., 2021).

3.2.3 AR(1) benchmark

Benchmark month-end mappings are produced by a univariate AR(1) model estimated on the same window as the DL regression. For the level target,

$$EAI_m = \alpha + \phi EAI_{m-1} + u_m, \tag{11}$$

and for the change target,

$$\Delta EAI_m = \alpha + \phi \Delta EAI_{m-1} + u_m. \tag{12}$$

The AR(1) provides a minimal persistence benchmark. This choice is standard in macro-finance out-of-sample predictive evaluations (West, 2006; Faust et al., 2013) and is particularly relevant in small samples, where more parameterized models can suffer from estimation noise that degrades out-of-sample performance even when additional predictors have true population coefficients of zero (Clark and West, 2007; Hubrich and West, 2010).

3.2.4 Information set and timing

Month-end predictions for month m use only information available as of the end of month m (m|m). In particular, RT_m is computed using daily stress observations within month m (including any carry-forward fills governed by the thin-trading protocol) and does not use information from month $m+1$ or later. Similarly, the benchmark AR(1) uses only lagged realizations of the relevant target series. This timing convention aligns the exercise with a monitoring/early-warning interpretation (Andreou et al., 2013; Gomez-Zamudio and Ibarra, 2017): the objective is to assess whether a market-implied monthly stress summary is informative for contemporaneous monthly activity y_m at the month-end origin (as opposed to a pure $m \to m+1$ forecast). Although the target is indexed by month m, its realization is not assumed to be observed at the month-end origin (consistent with delayed release and revision in data-lagged settings).

3.3 Window discipline

The evaluation follows a three-window structure:

- **Burn-in window:** used to initialize estimation.
- **Validation window:** used to assess sensitivity of performance and to report intermediate out-of-sample results without using the final holdout.
- **Out-of-sample (OOS) window:** the final evaluation window used for headline comparisons.

Windows are pre-specified and pinned to an `AS-OF` date to ensure that results can be reproduced exactly from a fixed information set. Model parameters are estimated once using the burn-in window and then held fixed when generating month-end predictions for the validation and OOS windows (no recursive re-estimation). This fixed-coefficient design is intentional: it mimics a deployable monitoring rule calibrated once and avoids chasing estimation noise in a short post-reset overlap sample.

Window start/end dates are reported as calendar dates; in monthly evaluation, windows map to the corresponding month endpoints.

Table 3 reports the window definitions used in the empirical analysis.

3.4 Evaluation metrics

Month-end mapping accuracy is evaluated using standard loss measures and summary diagnostics. Let $\hat{y}_{m|m}$ denote the month-end prediction for month m produced using the

Table 3: Month-end evaluation window definitions (ALL sample).

Window	Start	End	N_{months}
Burn-in	2022-03-15	2023-12-31	22
Validation	2024-01-01	2024-12-31	12
Out-of-sample	2025-01-01	2026-02-14	14

information set through the end of month m (prior to observing the released value of y_m), and let $e_m = y_m - \hat{y}_{m|m}$ denote the prediction error. The primary metrics are:

- Root mean squared error (RMSE),
- Mean absolute error (MAE),
- Correlation between y_m and $\hat{y}_{m|m}$,
- R^2, reported cautiously as a descriptive statistic.

To ensure comparability across models, all accuracy statistics are computed on a shared set of months for which both models produce month-end predictions.

Finally, the interpretation of results emphasizes time variation in predictive content. Rolling correlations and windowed performance comparisons are used to assess whether predictive relationships are stable or episodic, consistent with a regime-sensitive "post-reset price discovery" interpretation (Chirinko et al., 2018; Medioli et al., 2022) rather than a presumption of constant gains.

4 Results

This section reports empirical properties of the daily fiscal stress signal constructed from post-restructuring Puerto Rico GO bond spreads and evaluates its real-time month-end *mapping/monitoring* content for Puerto Rico's monthly Economic Activity Index (EAI). The presentation follows the evaluation workflow in Section 3. I first document signal behavior and the thin-trading environment under which it is constructed. I then separate lead content from contemporaneous co-movement using restricted distributed-lag structures estimated on the burn-in overlap (the post-2022 overlap in which both the fiscal stress signal and monthly EAI are observed), consistent with standard mixed-frequency practice that emphasizes timing rather than purely contemporaneous association (Andreou et al., 2013; Ghysels et al., 2006). Finally, I report windowed performance for EAI levels and month-over-month changes under the pre-specified segmentation in Table 3.

A key limitation is that the post-reset burn-in overlap used to estimate the distributed-lag (DL) mapping is inherently short in this post-reset setting: the standardized daily stress signal is defined only after the 252-trading-day rolling standardization window becomes available, monthly aggregation then requires within-month availability, and the inclusion of lag structure further reduces usable months. As a result, for levels the overlap contains only $N = 9$ monthly observations once lag structure and availability constraints are imposed (and $N = 10$ for the contemporaneous-only specification). The results in Sections 4.2–4.4 should therefore be read as evidence about an early post-restructuring price-discovery episode and a deployable fixed-rule monitoring exercise, not as stable structural estimates.

Throughout, interpretation emphasizes potential regime sensitivity: any predictive content may be concentrated during early post-reset price discovery and may attenuate as trading and expectations stabilize (Chirinko et al., 2018; Fuertes and Kalotychou, 2007). The post-restructuring context and instrument reset underpinning the sample definition are consistent with the documented GO settlement timeline and restructuring environment (Medioli et al., 2022; U.S. Government Accountability Office, 2025).

4.1 Signal validity and data-quality environment

Figure 5 plots the daily composite fiscal spread (constructed under the thin-trading protocol, including bounded carry-forward) and the standardized stress index used in the monthly regressions. The standardized series exhibits persistent movements over the overlap period, indicating a non-degenerate market-implied stress signal rather than purely transitory microstructure noise. This interpretation is consistent with evidence that properly measured credit spreads can act as early aggregators of information about real activity and incipient downturn risk (Faust et al., 2013; Andreou et al., 2013).

Because the municipal market trades infrequently, signal interpretation depends on market coverage (Green et al., 2007; Craig et al., 2018; Chalmers et al., 2019). Figure 6 reports the daily coverage ratio (fraction of configured bonds trading), and Figure 7 reports rolling no-trade and stale rates. These diagnostics confirm that the series is constructed in a thin-trading environment and motivate the explicit distinction between valid multi-bond days, stale one-bond days, and no-trade days. More broadly, nonsynchronous trading and stale pricing induce delayed incorporation of information and serial dependence in observed returns or spreads, implying that explicit protocols are required to interpret high-frequency signals extracted from illiquid markets (Lo and MacKinlay, 1989; Getmansky et al., 2004; Choi et al., 2021).

Finally, Table 4 assesses whether results are driven by a single security by estimating the baseline distributed-lag specification separately for each maturity-specific spread and for the median composite. Fit and coefficient magnitudes are broadly similar across CUSIPs and the composite, supporting the median aggregation as a robust summary of locally priced credit stress. Table 5 reports coefficient diagnostics for the baseline DL$(0, 1)$ specification estimated on the burn-in overlap; given the short overlap sample ($N = 9$), these estimates are treated as descriptive evidence of economically interpretable post-reset pricing rather than stable structural parameters. Accordingly, conventional small-sample inference (e.g., t-tests and p-values) should be viewed as diagnostic rather than dispositive.

To further assess whether the overlap association could plausibly reflect chance alignment or a single influential observation, I report two additional diagnostics. First, leave-one-out re-estimation shows that excluding any single month leaves the burn-in overlap fit within a narrow range ($R^2 \in [0.653, 0.805]$), indicating that the baseline overlap relationship is not driven by an individual month. Second, a permutation diagnostic based on 500 random reassignments of the fiscal-stress regressor yields a null R^2 distribution in which the observed overlap fit ($R^2 = 0.715$) exceeds the 95th percentile (0.661), with $p = 0.028$. Taken together with the cross-CUSIP consistency in Table 4, these checks support the interpretation that the burn-in overlap fit reflects a genuine early post-reset information episode, while remaining consistent with the broader evidence of regime sensitivity outside the overlap.

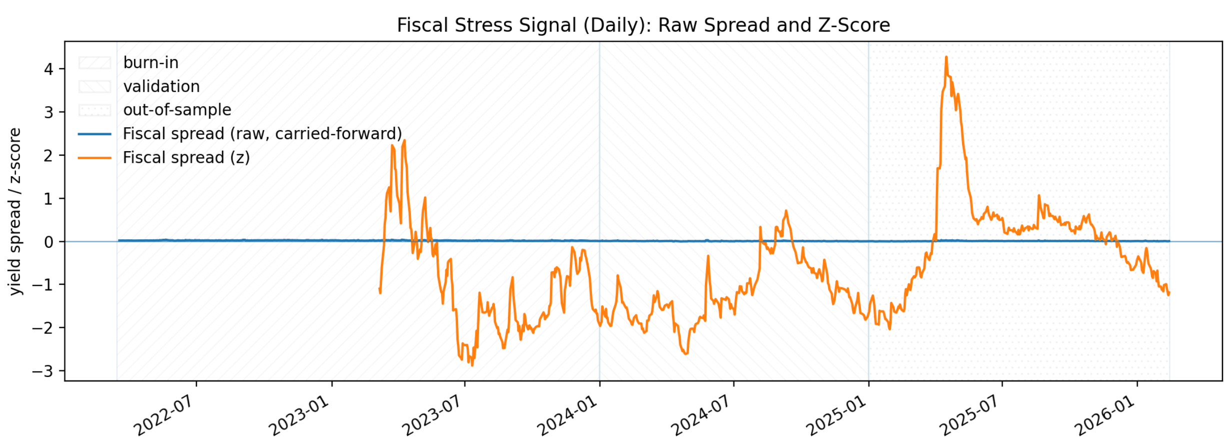

Figure 5: **Fiscal stress signal (daily): raw composite spread and standardized stress.** Sample begins post-settlement on 2022-03-15. The raw spread series is carried forward on stale/no-trade days under the thin-trading protocol; the standardized series is the stress index used in the monthly regressions. Shaded regions indicate burn-in, validation, and out-of-sample windows.

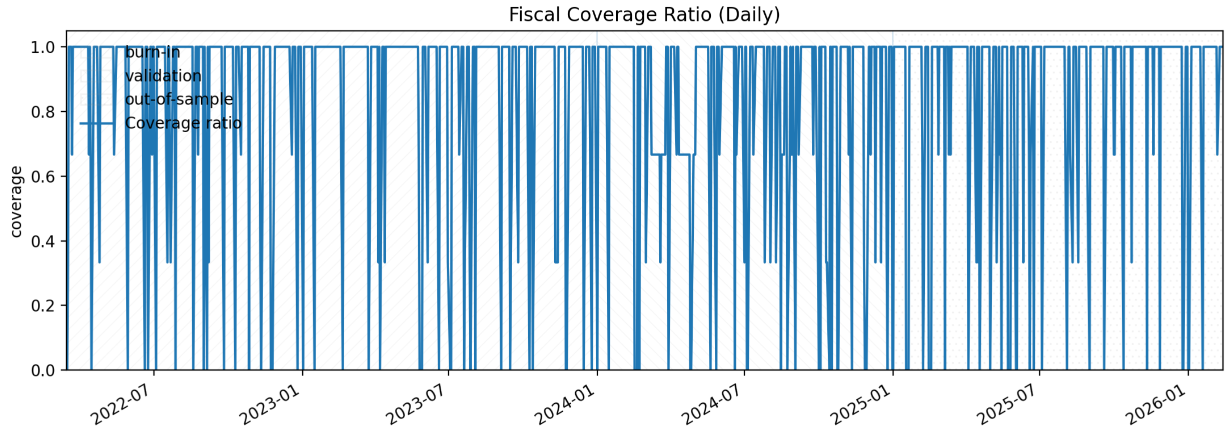

Figure 6: **Fiscal coverage ratio (daily).** Sample begins post-settlement on 2022-03-15. Daily coverage ratio under the multi-bond construction. Shaded regions indicate burn-in, validation, and out-of-sample windows.

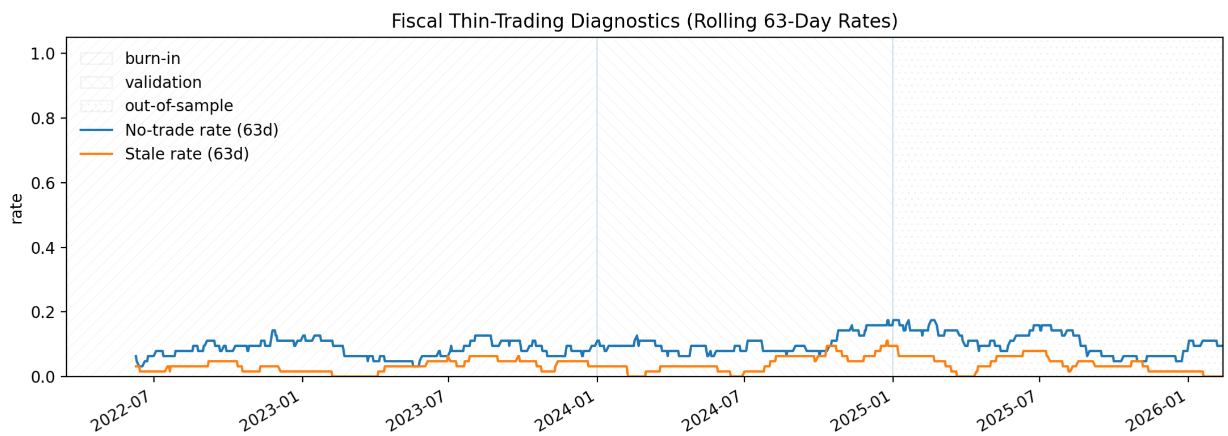

Figure 7: **Fiscal thin-trading diagnostics (rolling 63-day rates).** Sample begins post-settlement on 2022-03-15. Rolling no-trade and stale rates for the fiscal signal. Shaded regions indicate burn-in, validation, and out-of-sample windows.

Table 4: **Cross-bond consistency (burn-in overlap): composite vs single-CUSIP spreads.** Baseline DL$(0, 1)$ estimates for the median composite and for each maturity-specific spread. Similar fit and coefficient magnitudes across maturities support the composite as a robust summary.

Signal	CUSIP	N	R^2	RMSE	β_0	β_1
Median (3-bond)	All	9	0.715	0.429	0.607	-0.957
2035 spread (DGS7)	74514L3L9	9	0.732	0.416	0.440	-0.864
2037 spread (DGS10)	74514L3M7	9	0.759	0.395	0.417	-0.906
2046 spread (DGS20)	74514L3P0	9	0.516	0.559	0.418	-0.947

Table 5: **Coefficient diagnostics for the baseline DL specification (lags 0–1; burn-in overlap).** Reported are coefficient estimates with standard errors, t-statistics, and p-values for EAI levels and month-over-month changes (ΔEAI). R^2 and N refer to the burn-in overlap sample.

Target	Term	Estimate	Std. Err.	t-stat	p-value
Panel A: EAI level					
Level	Const	127.745	0.281	454.067	7.77×10^{-15}
Level	RT_m (β_0)	0.607	0.320	1.898	0.106
Level	RT_{m-1} (β_1)	-0.957	0.269	-3.556	0.012
Level	R^2	0.715			
Level	N	9			
Panel B: ΔEAI					
Delta	Const	0.244	0.270	0.903	0.401
Delta	RT_m (β_0)	0.157	0.307	0.512	0.627
Delta	RT_{m-1} (β_1)	-0.179	0.259	-0.693	0.514
Delta	R^2	0.075			
Delta	N	9			

4.2 Lead content versus contemporaneous co-movement

To separate lead content from within-month correlation, I compare three restricted specifications estimated on the burn-in overlap: (i) a contemporaneous-only regression, (ii) a lag-only regression, and (iii) the baseline DL$(0, 1)$ model. This separation mirrors the core identification concern in mixed-frequency monitoring/nowcasting with high-frequency financial indicators: contemporaneous association can reflect within-period co-movement, whereas lagged terms more directly capture lead content (Andreou et al., 2013; Gomez-Zamudio and Ibarra, 2017). Sample sizes differ because including a lagged term requires dropping the first available month in the overlap; hence Models B/C use one fewer observation than the contemporaneous-only specification. Table 6 reports fit metrics and coefficient estimates. The lag-only specification delivers materially higher R^2 and lower RMSE than the contemporaneous-only model, suggesting that the fiscal stress index contains short-horizon lead content rather than only contemporaneous co-movement.

The DL$(0, 1)$ specification combines both components and yields the best in-sample fit, but it also exhibits opposite-signed coefficients ($\beta_0 > 0$, $\beta_1 < 0$). Mechanically, this sign pattern implies that the mapping loads more heavily on an "innovation" or change-like component of the stress index (roughly $RT_m - \lambda RT_{m-1}$) than on the level of stress alone. In a thin market where transaction-implied spreads can jump with information arrival and episodic liquidity, such a structure is consistent with an impulse-and-partial-reversal pattern: the contemporaneous term captures the immediate stress signal, while the lagged term removes slow-moving persistence that can be induced by aggregation and bounded carry-forward. Given the short overlap sample, the DL coefficients are treated as descriptive evidence of timing and lead content in the early post-reset regime rather than as a stable structural parameterization.

Table 6: **Contemporaneous vs lagged vs distributed-lag structure (burn-in overlap).** Restricted-model comparison using standardized fiscal stress (RT_m). Model A uses only the contemporaneous term (m), Model B uses only the lagged term ($m-1$), and Model C uses both lags $(0, 1)$.

Model	Specification	N	R^2	RMSE	β_0	β_1
A	RT_m only (contemporaneous)	10	0.284	0.771	-0.475	
B	RT_{m-1} only (lagged)	9	0.544	0.543		-0.550
C	DL$(0, 1)$: RT_m and RT_{m-1}	9	0.715	0.429	0.607	-0.957

4.3 Real-time mapping test: EAI level (fiscal-only and composite, with AR(1) benchmark)

I next evaluate out-of-sample performance under the fixed window discipline in Table 3. Monthly evaluation is conducted on month-end targets; the final window end is pinned to the `AS-OF` date, which can truncate the last partial month. Because EAI_m is not generally available at month-end in real time, the exercise should be interpreted as a month-end *mapping/monitoring* exercise: month m predictions use information through the end of month m (summarized in RT_m and RT_{m-1}) to infer the activity level for month m that will be released subsequently. Table 7 reports windowed performance for the real-time month-end mapping of EAI levels using the fiscal-only signal (RTEP-FISC) under the baseline distributed-lag (DL) mapping. Performance is strongest on the burn-in overlap and weakens in the validation and out-of-sample windows, consistent with time variation in the mapping from fiscal stress to activity.

A core practical question is whether the composite index (RTEP-ALL) improves performance and whether any gains are meaningful relative to a simple autoregressive benchmark. Table 8 therefore reports a direct DL versus AR(1) comparison for the composite (RTEP-ALL) evaluated on identical months within each window. In-sample, DL and AR(1) are comparable. In validation and out-of-sample windows, however, AR(1) substantially dominates DL for the composite: validation RMSE falls from 0.596 (DL) to 0.411 (AR1), and out-of-sample RMSE falls from 1.004 (DL) to 0.445 (AR1). This pattern reinforces the interpretation that the high burn-in fit reflects early post-reset price discovery rather than a time-invariant mapping. More generally, it is also consistent with the standard nested-model forecast-evaluation result that, in finite samples, estimation noise in the larger model can degrade out-of-sample MSPE even when additional predictors have limited incremental content (Clark and West, 2007; Hubrich and West, 2010; West, 2006). Taken together, the headline finding for levels is that the distributed-lag mapping does not deliver stable improvements in point month-end mappings/nowcasts beyond simple autoregressive dynamics outside the burn-in overlap, even when the auxiliary equity component is included in the composite.

Table 7: **EAI level: DL performance using RTEP-FISC (fiscal-only).** Windowed evaluation over burn-in, validation, and out-of-sample (OOS) segments. Reported are R^2, RMSE, MAE, correlation, and the number of monthly observations (N).

Window	N	R^2	RMSE	MAE	Corr
Burn-in	9	0.7149	0.4294	0.3790	0.8455
Validation	12	0.1740	0.5882	0.4650	0.4673
OOS	11	-4.1905	0.9968	0.8053	0.2987

Table 8: **Composite (RTEP-ALL), EAI level: DL vs AR(1) benchmark on identical months.** Windowed evaluation for the composite index using DL$(0, 1)$ versus an AR(1) benchmark. Each window reports metrics computed on the same set of months for both models.

		DL			AR(1)		
Window	N	R^2	RMSE	MAE	R^2	RMSE	MAE
Burn-in	9	0.732	0.417	0.365	0.718	0.427	0.335
Validation	12	0.151	0.596	0.473	0.597	0.411	0.359
OOS	11	-4.262	1.004	0.816	-0.035	0.445	0.382

4.4 Real-time mapping test: ΔEAI

Table 9 reports the corresponding windowed results for the real-time month-end mapping of month-over-month changes in economic activity using the fiscal-only signal (RTEP-FISC) under the baseline DL mapping. Relative to the level target, the fiscal stress signal exhibits weaker and less stable mapping to ΔEAI_m across windows. This is consistent with the interpretation that the stress index primarily reflects slower-moving level dynamics (and persistence in broader conditions) rather than high-frequency month-to-month fluctuations, a pattern commonly observed when credit spread indicators are used for monitoring or nowcasting lower-frequency real activity measures (Faust et al., 2013; Andreou et al., 2013).

Table 9: **ΔEAI: DL performance using RTEP-FISC (fiscal-only).** Windowed evaluation over burn-in, validation, and out-of-sample (OOS) segments. Reported are R^2, RMSE, MAE, correlation, and the number of monthly observations (N).

Window	N	R^2	RMSE	MAE	Corr
Burn-in	9	0.0747	0.4128	0.3341	0.2733
Validation	12	-1.7359	0.5034	0.4548	0.2484
OOS	11	-0.2878	0.5147	0.3916	-0.0004

4.5 Regime sensitivity diagnostics for the composite

Two complementary diagnostics highlight instability in the composite construction and its relationship with activity. First, the expanding-window composite weight path (Table 10) shows that the calibrated equity weight varies over time and that the in-window explanatory content of both components declines as the estimation window expands. The equity weight is generally small relative to the fiscal weight, but rises late in the sample even as the component-by-component R^2 values fall, consistent with an unstable and weakening mapping rather than a stable incremental signal.

Second, the rolling 24-month R^2 diagnostic for standardized monthly EAI and standardized monthly signals (Table 11) indicates that, when defined for the composite, the rolling R^2 is modest (mean 0.064 over 2025-02 through 2025-11, max 0.146). This reinforces the interpretation that any strong in-sample association in the short burn-in overlap does not persist as a stable medium-run relationship.

Table 10: **Expanding-window composite weight path: snapshot.** Selected expanding-window estimates of composite weights and component R^2 diagnostics from the composite calibration routine.

Train end	N	w_{eq}	w_{fisc}	R^2_{eq}	R^2_{fisc}
2023-08-31	6	0.161	0.839	0.091	0.474
2024-10-31	20	0.115	0.885	0.039	0.301
2025-11-30	33	0.229	0.771	0.026	0.088

4.6 Incremental value of equity (one-shot comparison)

Finally, I assess whether adding the auxiliary equity component materially changes month-end mapping accuracy beyond the fiscal-only signal at the monthly horizon considered here. Figure 8 compares out-of-sample RMSE for the composite (RTEP-ALL) versus the fiscal-only variant (RTEP-FISC) under the distributed-lag specification. The comparison indicates no

Table 11: **Rolling 24-month R^2 summary (standardized monthly series).** Summary of rolling 24-month R^2 values computed as rolling correlation^2 between monthly z-scored EAI and monthly z-scored signals (standardized on the ALL burn-in window). The composite rolling R^2 is only defined once the composite series has sufficient overlap.

Series	Availability	Mean R^2	Min R^2	Max R^2
RT-EQ vs EAI	2022-03 to 2026-02 (45 months with defined 24m window)	0.053	0.000	0.284
RT-ALL vs EAI	2025-02 to 2025-11 (10 months with defined 24m window)	0.064	0.010	0.146

material improvement from adding equity at the monthly horizon. Combined with Table 8, the results support a conservative conclusion: while the fiscal signal exhibits interpretable movements and a lag-tilted mapping in the burn-in overlap, stable improvements in point month-end mappings/nowcasts beyond simple autoregressive dynamics are not present in validation and out-of-sample windows.

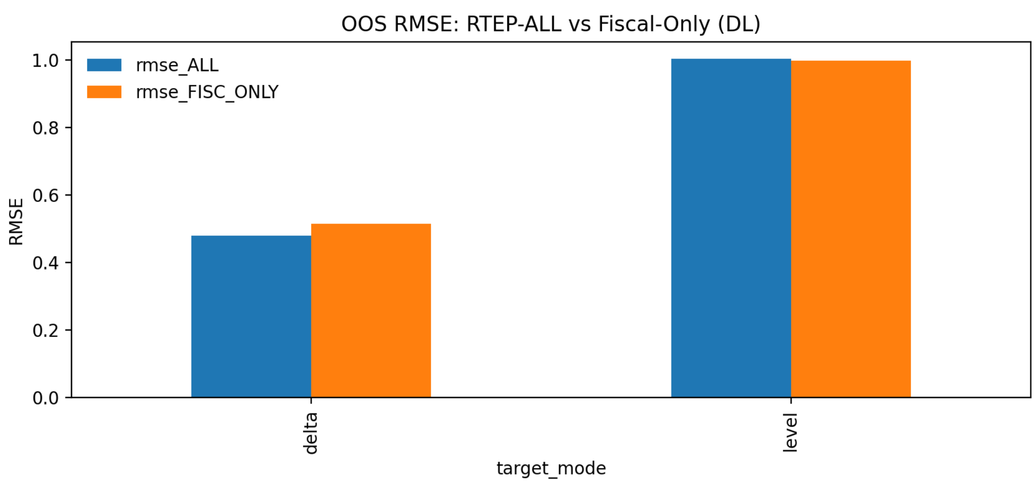

Figure 8: **Out-of-sample RMSE: RTEP-ALL vs fiscal-only (DL).** Lower values indicate better performance. The comparison is reported for both target modes (EAI level and ΔEAI).

5 Discussion

This section interprets the empirical patterns documented in Section 4. The key finding is episodic real-time monitoring/mapping content (not stable nowcast dominance): the fiscal stress signal shows economically interpretable lead content in the early post-restructuring overlap, while out-of-sample month-end mapping performance deteriorates as the sample advances. This pattern is consistent with a regime-sensitive "post-reset price discovery" channel rather than a time-invariant mapping from spreads to monthly activity.

5.1 What the burn-in estimates do (and do not) imply

The burn-in overlap results show that lagged fiscal stress is more informative than contemporaneous stress in restricted specifications, and that this relationship is broadly consistent across the three CUSIPs and the composite. Interpreted narrowly, these patterns support the claim that post-reset transaction-implied spreads embed forward-looking information about near-term economic conditions and can contribute to month-end monitoring (m|m) in data-lagged environments. Interpreted cautiously, they do not establish a stable structural relationship. The fiscal signal begins in 2022-03-15, and the distributed-lag mapping is therefore estimated on a necessarily short monthly overlap once lag structure and missingness are imposed. With limited overlap, coefficient estimates are best viewed as a descriptive calibration of the early post-restructuring regime rather than a parameterization expected to remain invariant. The purpose of the out-of-sample exercise is correspondingly narrow: to test whether the sign and direction of the relationship persist and whether the signal exhibits any incremental month-end mapping content relative to persistence benchmarks, not to claim durable dominance in point forecasts.

5.2 Why out-of-sample performance degrades

The out-of-sample results indicate that the DL specification does not reliably outperform the AR(1) benchmark for either target, and performance is weakest in the final holdout. Three non-exclusive mechanisms are consistent with this deterioration.

First, the post-restructuring sample is short. With limited monthly observations, estimation noise can dominate incremental predictors and degrade out-of-sample performance even when a predictor exhibits burn-in overlap fit (Clark and West, 2007; Hubrich and West, 2010; West, 2006).

Second, liquidity and trading intensity vary over time. Because the fiscal signal is constructed from transactions in a thin market, the effective information content of the daily series depends on coverage and the frequency of valid multi-bond trading days. Periods with fewer informative prints can reduce the quality of the monthly aggregation even when the construction protocol avoids uncontrolled interpolation.

Third, the mapping from spreads to activity may change as the market transitions from initial price discovery to a more stable expectation regime. Immediately after a restructuring, spreads may incorporate rapid learning about settlement mechanics, oversight constraints, and fiscal capacity. As those uncertainties resolve, spreads may move more with broader interest-rate conditions and risk appetite than with local near-term activity, weakening the relationship to the EAI.

5.3 Early-warning signal versus point-nowcast improvement

A central implication of the results is that "useful signal" need not imply "lower RMSE than an autoregressive benchmark." The fiscal stress index may still serve as an early-warning or monitoring device if it provides directional or state information—e.g., elevated stress identifying periods in which downside risks to activity are higher—even if the monthly DL mapping does not deliver consistent improvements in point nowcasts (Fuertes and Kalotychou, 2007; González-Rivera et al., 2019). This distinction is particularly relevant in data-lagged environments where the operational value of market indicators can lie in rapid detection of regime shifts rather than marginal reductions in average nowcast loss.

5.4 Portability to other settings

The construction and evaluation framework is portable to other small, data-lagged economies and sub-sovereign issuers, particularly following fiscal resets (restructuring, oversight transitions, or capital-structure revisions) that reopen price discovery. The key requirements are (i) locally relevant securities with observable transaction data, (ii) a deep benchmark curve for maturity-matched spread extraction, and (iii) explicit thin-trading rules with coverage diagnostics to avoid implicit interpolation. Where these conditions hold, transaction-implied credit spreads can provide a transparent complement to delayed macro releases, with the caveat that predictive content may be episodic and should be evaluated under strict window discipline and benchmark comparisons.

6 Conclusion

This paper asks whether transaction-implied credit spreads from post-restructuring Puerto Rico general obligation (GO) bonds provide a useful early-warning or month-end monitoring indicator, and whether they contain incremental real-time (m|m) monitoring/mapping signal for Puerto Rico's monthly Economic Activity Index (EAI). The setting is a "credit reset" episode following the March 15, 2022 restructuring settlement, in which refreshed trading in restructured GO securities can concentrate information into market spreads and potentially improve real-time monitoring in a data-lagged environment.

The paper's first contribution is the construction of a transparent, transaction-based fiscal stress index from locally relevant securities. The index is built from EMMA transaction-implied yields for three restructured GO maturities, benchmarked to maturity-matched U.S. Treasury yields, aggregated via a median composite, and standardized using a rolling procedure without leakage. Because municipal markets trade infrequently, the construction explicitly treats thin trading through a documented protocol (no-trade and stale-day flags, bounded carry-forward) and reports coverage diagnostics. The result is an auditable series designed for implementation in small, enclosed, or administratively constrained economies where timely macro indicators are limited.

The second contribution is an evaluation framework that emphasizes reproducibility and conservative inference. Month-end mappings (m|m) are generated under a strict information-set timing convention and pre-specified burn-in, validation, and out-of-sample windows, with model parameters estimated on the burn-in window and held fixed thereafter (no recursive re-estimation). Performance is assessed relative to a minimal AR(1) benchmark and reported in windowed form to allow for regime dependence rather than assuming constant gains.

Empirically, the fiscal stress index exhibits economically interpretable within-sample associations with EAI, and restricted distributed-lag comparisons indicate that lagged stress contains more information than purely contemporaneous co-movement in the burn-in overlap. Out of sample, however, the distributed-lag mapping does not consistently dominate the univariate benchmark for either EAI levels or month-over-month changes. Windowed performance patterns and diagnostics are consistent with regime-sensitive content: spreads appear most informative during early post-reset price discovery and less informative as trading, expectations, and broader financial conditions stabilize. Adding an auxiliary equity component does not materially improve performance at the monthly horizon considered. In this setting, the spreads' primary value is state detection and episode monitoring rather than average-error minimization in point predictions.

Taken together, the results support a conservative conclusion. Post-restructuring GO spreads can provide a transparent, market-implied monitoring indicator that may be useful for early warning while the official EAI is pending release, but they do not deliver stable improvements in point month-end mapping/nowcast accuracy relative to a persistence benchmark in the available sample. This distinction is operationally relevant in data-lagged settings: the value of market indicators may lie in identifying episodes of elevated stress and downside risk rather than in reducing average loss in point predictions.

Several extensions are natural. First, performance should be revisited as a longer post-reset sample accumulates, allowing more reliable assessment of stability and regime dependence. Second, state-space or time-varying parameter approaches could formalize the regime-sensitive interpretation by allowing the mapping from spreads to activity to evolve over time. Third, the framework can be applied to alternative macro targets (e.g., revenue collections, employment proxies, sectoral indicators) for which the informational advantage of high-frequency market data may be stronger. More broadly, the construction and evaluation protocol is portable to other sub-sovereigns and small economies following fiscal resets, provided that transaction data and benchmark yields are available and thin trading is handled explicitly.

A Data Description and Windows

Replication note. The replication package excludes raw EMMA trade-level extracts due to licensing and redistribution constraints on underlying transaction records. It includes the derived daily signal series used in estimation, monthly aggregates and evaluation outputs, code and configuration files, and cryptographic hashes that uniquely identify the raw inputs used to generate the reported results.

A.1 Data Summary

Table 12 provides an overview of all data series used in the analysis, including sample sizes, frequencies, and thin-trading coverage statistics.

Table 12: Data summary: all series used in the analysis.

Series	Frequency	Sample start–end	N	Coverage / thin-trading statistics
PR GO 2035 (74514L3L9)	Daily (trades)	2022-03-15–2026-02-13	1024	13.48% zero-trade; median run 1.0d; max 4d
PR GO 2037 (74514L3M7)	Daily (trades)	2022-03-15–2026-02-13	1024	14.55% zero-trade; median run 1.0d; max 15d
PR GO 2046 (74514L3P0)	Daily (trades)	2022-03-15–2026-02-13	1024	13.67% zero-trade; median run 1.0d; max 12d
UST DGS7/DGS10/DGS20	Daily	2010-01-01–2026-02-12	4205	—
EAI, ΔEAI	Monthly	1999-07-31–2025-11-30	317	—

Note: The ΔEAI series has $N = 316$ due to first-differencing.

EMMA extraction workflow (manual, without redistributing raw extracts). EMMA transaction data were collected manually from the EMMA web interface for each configured PR GO

CUSIP. For each CUSIP, I navigated to the issue security page and the trade activity view, set the date range to 2022–2026, and filtered the transaction type to *Inter-dealer trade*. Because EMMA does not provide a bulk export pathway for this workflow, the trade table was paginated; I copied and pasted each page of transactional rows into an Excel workbook (on the order of dozens of pages per CUSIP) and then consolidated and cleaned those pages using a deterministic parser. The parser scans each sheet for the EMMA trade-table header (anchored on `Trade Date/Time`), reads the table body, concatenates sheets, and then filters to `Trade Type == "Inter-dealer trade"` while dropping non-row artifacts (e.g., header repeats and footer text) that appear in the pasted pages. The resulting consolidated inter-dealer trade table (CSV/XLSX) is the input used by the downstream yield aggregation and signal-construction pipeline.

A.2 Window Definitions

Table 13 reports the exact window boundaries used for burn-in, validation, and out-of-sample evaluation.

Table 13: Month-end evaluation window definitions.

Window	Start	End	N_{months}
Burn-in	2022-03-15	2023-12-31	22
Validation	2024-01-01	2024-12-31	12
Out-of-sample	2025-01-01	2026-02-14	14

Note: Monthly regression sample sizes are smaller than the window lengths shown in Table 13 because they are restricted to months with overlapping EAI and aggregated RT-FISC observations. In particular, the burn-in coefficient diagnostics in Appendix D are based on the burn-in *overlap* sample with $N = 9$ months.

B Thin-Trading Protocol Summary

Table 14 summarizes thin-trading diagnostics for the fiscal signal. These statistics validate the carry-forward discipline and coverage properties described in Section 2.3.

Table 14: Thin-trading diagnostics for the fiscal signal (daily; EMMA trades).

Diagnostic	Value	Notes
Zero-trade days (%)	9.47	No configured bonds trade
Stale days (%)	3.52	Only one bond trades (multi-bond mode)
Median carry duration (days)	1.0	Consecutive carry-forward length (applied)
Max consecutive carry (days)	4	Capped at MAX_FISCAL_CARRY_DAYS
Mean coverage ratio	0.861	Average fraction trading

C Replication Package Contents

The replication package contains the derived time series and outputs required to reproduce all tables and figures in this paper without redistribution of raw EMMA trade-level extracts. Included items are:

- **Derived daily fiscal signal and diagnostics.**
  - `fisc_flags_daily.csv` (N=1,024 days; 2022-03-15 to 2026-02-14): Daily thin-trading flags and diagnostics including `n_traded`, `coverage_ratio`, `is_stale`, `is_no_trade`, `is_carry_event`, and `single_cusip_mode`.
  - `fisc_spread_raw_daily.csv` (N=1,024 days): Daily composite fiscal spread (GO yield minus maturity-matched Treasury) after applying the documented filtering and thin-trading protocol.
  - `fisc_spread_z_daily.csv` (N=1,024 days): Standardized daily fiscal spread used in all distributed-lag regressions.
- **Monthly aggregates and evaluation outputs.**
  - `monthly_fisc_z.csv`: Monthly aggregation of the standardized fiscal stress series aligned to the EAI calendar.
  - `evaluation_fisc_only_results.csv` and `evaluation_fisc_only_delta_results.csv`: Windowed evaluation outputs for RTEP-FISC (levels and ΔEAI).
  - `evaluation_all_results.csv` and `evaluation_all_delta_results.csv`: Windowed evaluation outputs for RTEP-ALL (levels and ΔEAI).
- **Configuration and reproducibility metadata.**
  - `metadata.json`: Full configuration snapshot (windows, CUSIP mappings, Treasury assignments, protocol parameters, and software environment).

  - `input_hashes.json`: SHA-256 checksums uniquely identifying the raw EMMA and FRED input files used in the pipeline.

- **Robustness outputs.**
  - `leave_one_out_results.csv`: Leave-one-out diagnostics for burn-in regressions.
  - `permutation_r2_distribution.csv` and `permutation_test_summary.json`: Permutation-test outputs for the burn-in R^2 diagnostic.

These materials enable exact replication of the estimation, aggregation, evaluation, and figure generation steps reported in the paper using the derived series, while respecting redistribution constraints on raw transaction records.

D Robustness and Stability Battery

D.1 Coefficient Diagnostics

This subsection reproduces the coefficient diagnostics reported in the main text (Table 5) for archival transparency. Table 15 reports coefficient estimates, standard errors, t-statistics, and p-values for the baseline DL(0,1) specification on both EAI levels and month-over-month changes.

Table 15: Coefficient diagnostics for the baseline DL specification (lags 0–1; burn-in overlap).

Target	Term	Estimate	Std. Err.	t-stat	p-value
Panel A: EAI level					
Level	Const	127.745	0.281	454.067	7.77×10^{-15}
Level	FISC_z_lag0	0.607	0.320	1.898	0.106
Level	FISC_z_lag1	−0.957	0.269	−3.556	0.012
Level	R^2	0.715			
Level	N	9			
Panel B: ΔEAI					
Delta	Const	0.244	0.270	0.903	0.401
Delta	FISC_z_lag0	0.157	0.307	0.512	0.627
Delta	FISC_z_lag1	−0.179	0.259	−0.693	0.514
Delta	R^2	0.075			
Delta	N	9			

D.2 Cross-Bond Consistency

Table 16 reports DL(0, 1) diagnostics estimated separately for each maturity-specific CUSIP spread and for the median composite. Similar fit and coefficient magnitudes across maturities indicate that the fiscal signal is not driven by a single security. This table reproduces the cross-bond consistency results reported in the main text (Table 4) for completeness.

Table 16: Cross-bond consistency: composite vs single-CUSIP fiscal spreads (burn-in overlap).

Signal	CUSIP	N	R^2	RMSE	β_0	β_1
Median (3-bond)	All	9	0.715	0.429	0.607	−0.957
2035 spread (DGS7)	74514L3L9	9	0.732	0.416	0.440	−0.864
2037 spread (DGS10)	74514L3M7	9	0.759	0.395	0.417	−0.906
2046 spread (DGS20)	74514L3P0	9	0.516	0.559	0.418	−0.947

D.3 Subsample Stability

Table 17 evaluates whether results are driven by a particular subsample period within the burn-in overlap. The overlap sample is split into first and second halves to assess temporal stability of coefficient estimates.

Table 17: Subsample stability: coefficient estimates across burn-in overlap subsamples.

Subsample	N	β_0	β_1	R^2	RMSE
Full burn-in overlap	9	0.607	−0.957	0.715	0.429
First half (2022-03 to 2023-07)	4	0.065	−0.160	0.981	0.013
Second half (2023-08 to 2023-12)	5	0.470	−0.335	0.175	0.324

Note: Subsample estimates are based on very small N (4 and 5 months, respectively) and are reported as a sensitivity check rather than as reliable standalone fits. The extremely low RMSE in the first-half subsample reflects the small sample size and should be interpreted with caution.

D.4 Leave-One-Out Stability

Figure 9 presents leave-one-out diagnostics showing the impact of each month on the burn-in overlap regression fit. The month-by-month stability check identifies whether any single observation drives the baseline results.

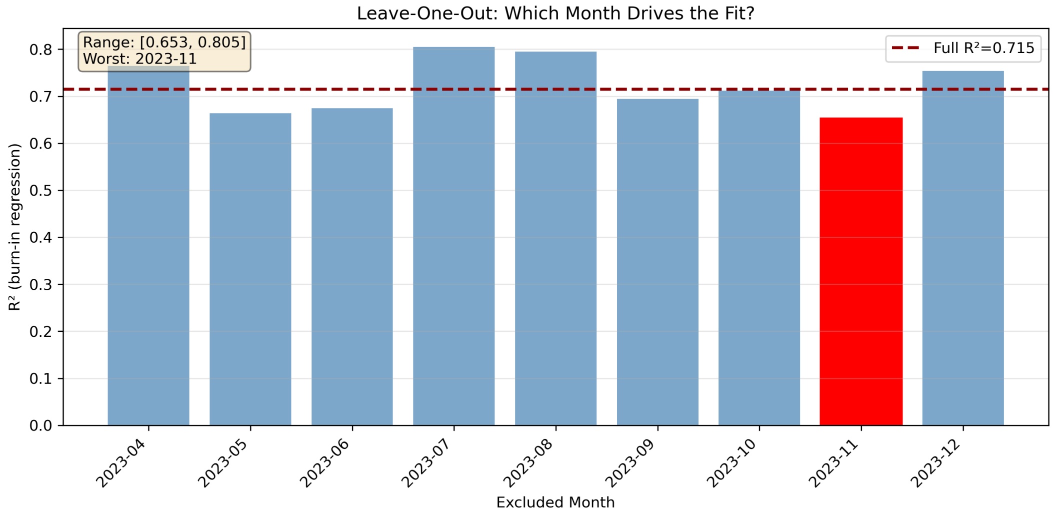

Figure 9: Leave-one-out stability analysis. Each bar shows burn-in overlap R^2 when excluding one month. Range: [0.653, 0.805]; worst exclusion: 2023-11.

D.5 Permutation Test

Figure 10 shows the permutation test distribution for the burn-in overlap R^2. Under the null hypothesis of no relationship between fiscal stress and EAI, the fiscal stress regressor is randomly permuted 500 times and the DL model is re-estimated. The observed $R^2 = 0.715$ exceeds the 95th percentile of the null distribution (0.661), yielding $p = 0.028$.

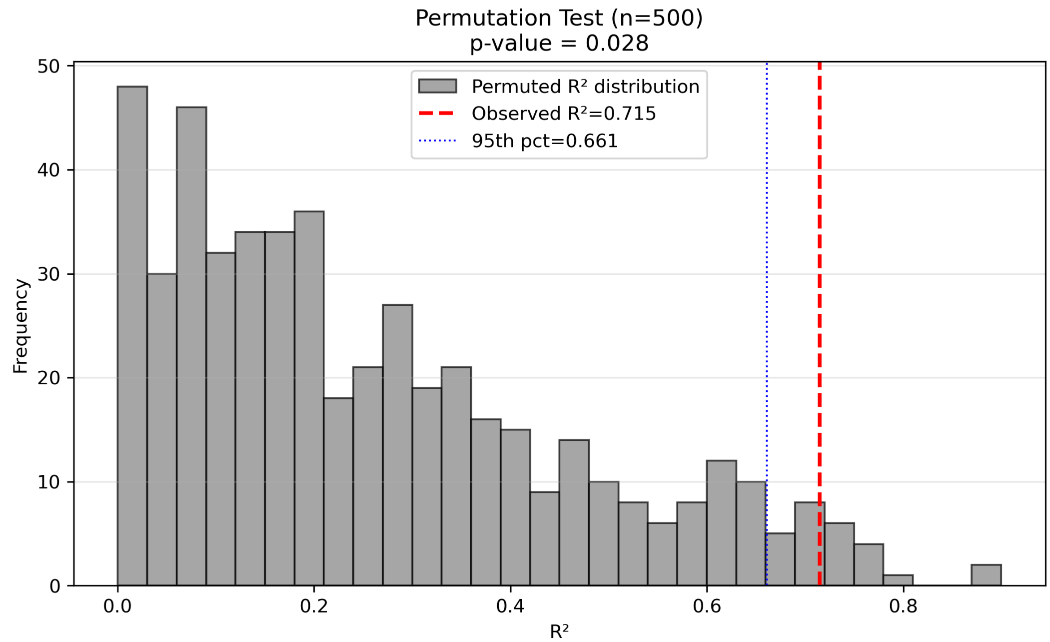

Figure 10: Permutation test distribution (n=500). The observed $R^2 = 0.715$ exceeds the 95th percentile ($p = 0.028$), providing evidence that the burn-in overlap association exceeds chance under this diagnostic.

E Nowcast Evaluation: Complete Results

This section provides complete windowed evaluation metrics for (i) the fiscal-only specification and (ii) the composite specification used in the incremental comparison in Figure 8. The fiscal-only metrics correspond to the outputs underlying Tables 7 and 9 in the main text.

E.1 Fiscal-Only Specification

Table 18 provides the complete evaluation metrics for the primary fiscal-only specification across all windows and target modes.

E.2 Composite Specification

Table 19 provides the complete evaluation metrics for the composite specification (RTEP-ALL) that combines fiscal and equity components. Results are reported for both target modes to support the incremental comparison in Figure 8.

Table 18: Complete evaluation metrics: fiscal-only DL specification (RTEP-FISC).

Window	Target	N	R^2	RMSE	MAE	Corr
Panel A: EAI level						
Burn-in	Level	9	0.715	0.429	0.379	0.846
Validation	Level	12	0.174	0.588	0.465	0.467
OOS	Level	11	−4.191	0.997	0.805	0.299
Panel B: ΔEAI						
Burn-in	Delta	9	0.075	0.413	0.334	0.273
Validation	Delta	12	−1.736	0.503	0.455	0.248
OOS	Delta	11	−0.288	0.515	0.392	−0.000

Table 19: Complete evaluation metrics: composite DL specification (RTEP-ALL).

Window	Target	N	R^2	RMSE	MAE	Corr
Panel A: EAI level						
Burn-in	Level	9	0.732	0.417	0.365	0.855
Validation	Level	12	0.151	0.596	0.473	0.455
OOS	Level	11	−4.262	1.004	0.816	0.306
Panel B: ΔEAI						
Burn-in	Delta	9	0.052	0.418	0.340	0.228
Validation	Delta	12	−1.687	0.499	0.447	0.272
OOS	Delta	11	−0.123	0.481	0.382	0.083

F Reproducibility Snapshot

The replication package includes a complete configuration snapshot used to generate all results. Key elements include:

- **AS-OF date:** 2026-02-14
- **Window definitions:** Burn-in, validation, and OOS boundaries
- **Thin-trading parameters:** `max_fiscal_carry_days = 20` (protocol cap)
- **CUSIP mappings:** 2035 → 74514L3L9, 2037 → 74514L3M7, 2046 → 74514L3P0
- **Treasury maturity assignments:** 2035 → DGS7, 2037 → DGS10, 2046 → DGS20
- **Input data hashes:** SHA-256 checksums uniquely identifying the raw EMMA trades extract and FRED yield inputs used to construct the derived series
- **Software versions:** Python 3.x, pandas, numpy, statsmodels

Note: The AS-OF date (2026-02-14) refers to the pipeline pinning date for reproducibility. Actual data availability ends 2026-02-13 for EMMA trades and 2026-02-12 for FRED Treasury yields.

References

Andreou, E., Ghysels, E., and Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2):240–251.

Chalmers, J., Liu, Y. S., and Wang, Z. J. (2019). The difference a day makes: Timely disclosure and trading efficiency in the muni market. *SSRN Electronic Journal*.

Chari, A., Leary, R., and Phan, T. (2017). The costs of (sub) sovereign default risk: Evidence from puerto rico. Working Paper 24108, National Bureau of Economic Research.

Chirinko, R., Chiu, R., and Henderson, S. (2018). What went wrong?: The puerto rican debt crisis and the "treasury put". Working paper, University of Illinois at Chicago and CESifo.

Choi, J., Kronlund, M., and Oh, J. Y. J. (2021). Sitting bucks: Stale pricing in fixed income funds. Working paper, University of Illinois at Urbana-Champaign, Tulane University, and Hanyang University.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Craig, L., Kim, A., and Woo, S. W. (2018). Pre-trade information in the municipal bond market. White paper, U.S. Securities and Exchange Commission, Division of Economic and Risk Analysis (DERA).

Departamento de Hacienda de Puerto Rico (2023). Departamento de hacienda recauda $1,126.8 millones durante el mes de enero del año fiscal 2023. Comunicado de prensa, Gobierno de Puerto Rico.

Department of Economic Development and Commerce (2026). The puerto rico economic activity index (pr-eai): October & november 2025. Technical report, Government of Puerto Rico, San Juan, Puerto Rico.

Faust, J., Gilchrist, S., Wright, J. H., and Zakrajšek, E. (2013). Credit spreads as predictors of real-time economic activity: A bayesian model-averaging approach. *The Review of Economics and Statistics*, 95(5):1501–1519.

Fuertes, A.-M. and Kalotychou, E. (2007). Optimal design of early warning systems for sovereign debt crises. *International Journal of Forecasting*, 23(1):85–100.

Getmansky, M., Lo, A. W., and Makarov, I. (2004). An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3):529–609.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The midas touch: Mixed data sampling regression models. Scientific Series 2004s-20, CIRANO, Montréal.

Ghysels, E., Sinko, A., and Valkanov, R. (2006). Midas regressions: Further results and new directions. Technical report, University of North Carolina at Chapel Hill and UCSD. Working Paper.

Gomez-Zamudio, L. M. and Ibarra, R. (2017). Are daily financial data useful for forecasting gdp? evidence from mexico. Working Papers 2017-17, Banco de México.

González-Rivera, G., Maldonado, J., and Ruiz, E. (2019). Growth in stress. *International Journal of Forecasting*, 35(3):948–966.

Green, R. C., Hollifield, B., and Schürhoff, N. (2007). Financial intermediation and the costs of trading in an opaque market. *The Review of Financial Studies*, 20(2):275–314.

Hubrich, K. and West, K. D. (2010). Forecast evaluation of small nested model sets. *Journal of Applied Econometrics*, 25(4):574–594.

Lo, A. W. and MacKinlay, A. C. (1989). An econometric analysis of nonsynchronous trading. Working Paper 3003-89-EFA, Alfred P. Sloan School of Management, Massachusetts Institute of Technology.

Manasse, P., Roubini, N., and Schimmelpfennig, A. (2003). Predicting sovereign debt crises. IMF Working Paper WP/03/221, International Monetary Fund.

Medioli, A., Hampton, T., Chea, P., and Raimes, E. (2022). Puerto rico bondholder recovery patterns echo major municipal bankruptcies. Sector comment, Moody's Investors Service.

Qian, M. (2011). Stale prices and the performance evaluation of mutual funds. *Working Paper, National University of Singapore.*

U.S. Government Accountability Office (2025). U.s. territories: Public debt and economic outlook - 2025 update. Report to Congressional Committees GAO-25-107560, U.S. Government Accountability Office.

West, K. D. (2006). Forecast evaluation. In *Handbook of Economic Forecasting*, volume 1, pages 99–134. Elsevier.