Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 02-2026

Sentiment Utopia Index: Rhetorical Structure and Corpus-Relative Financial Sentiment

Stelios Arvanitis

January 2026

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

Sentiment Utopia Index: Rhetorical Structure and Corpus-Relative Financial Sentiment

Stelios Arvanitis*

January 25, 2026

Abstract

This paper introduces the Sentiment Utopia Index (SUI), a functional and corpus-relative measure of financial sentiment that explicitly incorporates rhetorical structure within texts. Each document is represented as a sentiment trajectory over normalized rhetorical time, and sentiment extremeness is defined as a weighted shortfall from corpus-level bullish and bearish envelopes constructed under real-time information constraints. Rhetorical emphasis is treated as a structured form of model uncertainty and explored through a finite, interpretable family of parametric weighting schemes. An empirical application illustrates the use of the SUI as a predictive signal for daily realized volatility in a heterogeneous autoregressive framework using SEC filings and an OHLC-based volatility proxy.

Keywords: Financial sentiment; Text-based indicators; Functional data analysis; Rhetorical structure; Narrative economics; Realized volatility forecasting; Corpus normalization; Dominance relations

*Athens University of Economics and Business, Greece, email: stelios@aueb.gr

1 Introduction

A growing strand of literature documents that textual and linguistic information may contain economically relevant signals about expectations, uncertainty, and risk beyond what is usually captured by traditional quantitative indicators. Early work by Tetlock (2007) showed that media tone contains predictive ability regarding market returns and trading behavior. Text-based measures capture economically meaningful variation in policy uncertainty (Baker et al., 2016), macroeconomic conditions (Kalamara et al., 2022), and financial volatility (Hassan et al., 2019; Manela and Moreira, 2017; Bybee et al., 2024). These contributions establish that narrative and linguistic information may not be merely descriptive but constitutes a distinct informational channel through which beliefs, expectations, and risk perceptions enter the dynamics of some economic and financial phenomena.

This paper proposes a conservative extension of standard text-based sentiment analysis that explicitly incorporates two features: first, financial texts are rhetorically structured: the position of statements within a document often reflects their informational role, emphasis, or intended salience. Second, the economic meaning of a given sentiment realization is inherently relative to the surrounding textual environment: a mildly negative tone may be informative in a tranquil narrative regime but quite unremarkable in a period of extreme pessimism. Existing indices rarely incorporate either rhetorical structure or corpus-relative normalization in a systematic way.

We represent each document as a functional sentiment profile over normalized rhetorical time, and we normalize each profile relative to corpus-level historical envelopes that define the feasible range of sentiment expressions. This construction yields a family of scalar indices-the Sentiment Utopia Indices (SUI)- that measure how close a document comes to the historically most bullish or most bearish rhetorical trajectories observed in the corpus, under a given weighting of rhetorical positions. The indices are interpretable, scale-free, and comparable over time by construction, while, they preserve information about rhetorical location that is lost in unstructured aggregation. Furthermore, rhetorical weighting is treated as a controlled sensitivity dimension, and simple parametric weight families are used in order to explore robustness.

We illustrate the empirical relevance of the construction in a forecasting application for daily realized volatility. Building on the heterogeneous autoregressive (HAR) framework of Corsi (2009), we examine whether the SUI contains incremental predictive information beyond standard volatility persistence components.

Relative to the existing literature, the contribution of this paper is methodological rather than algorithmic. We do not propose a new language model, a new embedding architecture, or a new forecasting paradigm. Instead, we propose a way of representing, normalizing, and ordering sentiment profiles emerging from language models, that (i) respects the rhetorical structure of financial texts, (ii) is explicitly relative to the historical narrative environment, and (iii) admits an order-theoretic interpretation in terms of dominance and extremeness. This perspective is complementary to existing approaches.

The remainder of the paper is organized as follows. Section 2 introduces the construction of functional sentiment profiles, corpus envelopes, rhetorical weighting, the Sentiment Utopia Index and discusses extensions and potential applications. Section 3 describes the data, the implementation specifics of the SUI methodology, the volatility proxy, and the forecasting design, and reports and discusses results regarding the aforementined predictive exercise. The final section concludes.

2 Methodology: constructing the SUI

This section defines the objects used throughout the paper and describes their construction in a way that applies to a general universe of temporally located financial documents. The empirical implementation in §3 instantiates these definitions using SEC EDGAR filings and a specific forecasting task, but the definitions below are not tied to that particular corpus or outcome variable.

2.1 Document universe, timing, and preprocessing

Let $\{d_t\}_{t \in \mathcal{T}}$ denote a sequence of financial documents observed over trading days τ (e.g., regulatory filings, earnings-call transcripts, news, press releases). Each document d_τ is timestamped and associated with an information set constraint: all quantities constructed for day t are required to be measurable with respect to information available by the end of day t. This real-time measurability condition is essential for predictive uses and rules out look-ahead normalization schemes. The available documents are assumed converted to plain text, stripped of markup and boilerplate, and segmented into sentences.

Issues regarding control of computational cost and/or of uniform representation across heterogeneous document lengths, the corpus of available documents actually used in the implementation may be further fine tuned via truncation rules that discard documents exceeding a given threshold of sentences' cardinality etc. Such restrictions are not inherent to the SUI construction but reflect the current implementation choices given the runtime budget.

2.2 Sentence-level sentiment scoring

Let document d_k contain n_k sentences $s_1^{(k)}, \ldots, s_{n_k}^{(k)}$. Each sentence receives a real-valued sentiment score

$$\sigma_i^{(k)} \in \mathbb{R}, \qquad i = 1, \ldots, n_k,$$

where $\sigma_i^{(k)} > 0$ indicates bullish tone, $\sigma_i^{(k)} < 0$ indicates bearish tone, and $\sigma_i^{(k)} = 0$ represents neutrality. The scoring function may be produced by a pretrained financial language model (e.g., FinBERT-see Araci (2019); Liu et al. (2021)) or an alternative estimator, e.g. DistilFinRoBERTa-see Sanh et al. (2019); the SUI construction below treats the resulting sentence scores as primitives.

In the implementation below, sentiment scores are produced by FinBERT at the sentence level and then treated as numerical observations along the rhetorical progression of each document.

2.3 Functional sentiment profiles over rhetorical time

The key methodological step is to represent each document as a function over standardized rhetorical time $x \in [0, 1]$, capturing how tone evolves from the beginning to the end of the document. A direct representation is the step-function profile

$$\sigma_k(x) := \sigma_i^{(k)}, \qquad x \in \Big(\frac{i-1}{n_k}, \frac{i}{n_k}\Big], \quad i = 1, \ldots, n_k, \tag{1}$$

which assigns each sentence score to its corresponding rhetorical interval.

For numerical stability and comparability across documents, we work with a discretized functional representation on a fixed grid of size G. Let x_g denote grid points corresponding to G equally spaced bins over $[0, 1]$, and let B_g be the set of sentences whose rhetorical time falls into bin g. The bin-mean profile is defined as

$$\sigma_k(x_g) := \frac{1}{|B_g|} \sum_{i \in B_g} \sigma_i^{(k)}, \qquad g = 1, \ldots, G, \tag{2}$$

with the convention that empty bins may be imputed by local interpolation or left missing (the implementation uses a binning procedure that yields a complete G-vector for each document under the imposed truncation rules).

In the empirical implementation, we set $G = 100$ and use the bin-mean representation (2).

2.4 Daily aggregation across multiple documents

When multiple documents arrive on the same day t, the methodology requires a rule that maps the set $\{d_{t,1}, \ldots, d_{t,m_t}\}$ to a single daily profile $\sigma_t(\cdot)$. In the current implementation we use a simple average-profile aggregation:

$$\sigma_t(x_g) := \frac{1}{m_t} \sum_{j=1}^{m_t} \sigma_{t,j}(x_g), \qquad g = 1, \ldots, G. \tag{3}$$

This choice preserves across document profiles linearity and produces a daily functional time series $\{\sigma_t(\cdot)\}$ aligned with trading days. Importantly, the SUI construction below applies equally to alternative aggregation rules-see the discussion in the penultimate paragraph of Section 2.10 below; we fix (3) to keep the empirical design minimal.

2.5 Corpus-relative envelopes under real-time constraints

The SUI measures extremeness relative to historically (or temporally local) observed bounds at each rhetorical position. Let $\sigma_t(x_g)$ denote the daily sentiment profile at day t on grid points x_g. Define the upper and lower envelopes (bullish and bearish bounds respectively) as functions of x. In an unrestricted retrospective setting, one could define global envelopes over the full sample by taking suprema and infima over all t. However, such fixed envelopes generally violate real-time measurability constraints in forecasting contexts because for example they incorporate future documents and/or they do not discard information that would not be present in rolling windows formalizations.

We therefore construct time-indexed envelopes using only information available up to day t. Two schemes are implemented.

Expanding envelopes. For each t and each grid point x_g,

$$\overline{\sigma}_t^{\exp}(x_g) := \max_{1 \leq s \leq t} \sigma_s(x_g), \qquad \underline{\sigma}_t^{\exp}(x_g) := \min_{1 \leq s \leq t} \sigma_s(x_g). \tag{4}$$

Expanding envelopes provide a real-time historical normalization frame that stabilizes as the sample grows.

Rolling envelopes. Fix a window length K (in trading days). For each t and each x_g,

$$\overline{\sigma}_t^{\text{roll}}(x_g) := \max_{t-K+1 \leq s \leq t} \sigma_s(x_g), \qquad \underline{\sigma}_t^{\text{roll}}(x_g) := \min_{t-K+1 \leq s \leq t} \sigma_s(x_g), \tag{5}$$

with the obvious modification when $t < K$. Rolling envelopes adapt to local regimes of communication and sentiment expression.

2.6 Rhetorical weighting: parametric family and grid

The SUI allows rhetorical emphasis to vary across a document via a nonnegative weight function $w(x)$ over rhetorical time $x \in [0, 1]$. In continuous form, one may normalize weights to integrate to one. In discretized approximations, normalization corresponds to scaling weights to sum to one over grid points.

We adopt a parametric exponential–quadratic family,

$$w_\theta(x) = \exp(\theta_1 x + \theta_2 x^2), \tag{6}$$

where $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ governs the shape of rhetorical emphasis. This parametric family generates a restricted but interpretable class of rhetorical weighting schemes. The sign of θ_2 governs the curvature of $\log w_\theta(x)$: when $\theta_2 < 0$ the weight function is log-concave and admits at most one interior mode, while $\theta_2 > 0$ yields a log-convex shape with mass concentrated toward the boundaries of the document. The linear term θ_1 interacts with θ_2 to shift the location of the dominant region of emphasis, determining whether weight is tilted toward earlier or later rhetorical positions. As a result, the family spans monotone emphasis patterns as well as single-mode interior emphasis, while remaining computationally simple and stable for large-scale sensitivity analysis.

In applied settings, the resulting index may be evaluated over a grid of parameter vectors θ within the exponential–quadratic family. This allows for systematic variation in rhetorical emphasis profiles (e.g., early, late, or non-monotonic emphasis), and supports robustness analysis with respect to the weighting specification. Grid-based evaluation also enhances numerical stability and enables sensitivity diagnostics across alternative parametrizations. The specific grid used in the empirical implementation is described in the following section.

2.7 Definition of the Sentiment Utopia Index (SUI)

We are now ready to provide the definition of SUI. Fix an envelope scheme $\mathsf{E} \in \{\exp, \text{roll}\}$ and a weight function $w(\cdot)$. Define the bullish shortfall functional as the weighted distance of the profile σ_t from the bullish envelope; the integrals below are interpreted in the Lebesgue-Stieljes sense:

$$\mathrm{SF}^{(+)}_{t,\mathsf{E}}(w) := \int_0^1 w(x)\,[\overline{\sigma}^{\mathsf{E}}_t(x) - \sigma_t(x)]_+\,dx, \tag{7}$$

where $[a]_+ = \max(a, 0)$. The corresponding maximal feasible shortfall under the same envelope is

$$\mathrm{MSF}^{(+)}_{t,\mathsf{E}}(w) := \int_0^1 w(x)\,[\overline{\sigma}^{\mathsf{E}}_t(x) - \underline{\sigma}^{\mathsf{E}}_t(x)]\,dx. \tag{8}$$

The bullish SUI is the normalized complement:

$$\mathrm{SUI}^{(+)}_{t,\mathsf{E}}(w) := 1 - \frac{\mathrm{SF}^{(+)}_{t,\mathsf{E}}(w)}{\mathrm{MSF}^{(+)}_{t,\mathsf{E}}(w)} \in [0, 1]. \tag{9}$$

Thus $\mathrm{SUI}^{(+)}_{t,\mathsf{E}}(w) = 1$ if the profile coincides with the bullish envelope almost everywhere (relative to w), and it is near zero when the profile is far from the bullish bound throughout rhetorical time.

Dually, define the bearish shortfall functional as the weighted distance of the profile from the bearish envelope:

$$\mathrm{SF}^{(-)}_{t,\mathsf{E}}(w) := \int_0^1 w(x)\,[\sigma_t(x) - \underline{\sigma}^{\mathsf{E}}_t(x)]_+\,dx, \tag{10}$$

with the same maximal range $\mathrm{MSF}^{(-)}_{t,\mathsf{E}}(w) := \mathrm{MSF}^{(+)}_{t,\mathsf{E}}(w)$. The bearish SUI is

$$\mathrm{SUI}^{(-)}_{t,\mathsf{E}}(w) := 1 - \frac{\mathrm{SF}^{(-)}_{t,\mathsf{E}}(w)}{\mathrm{MSF}^{(+)}_{t,\mathsf{E}}(w)} \in [0,1]. \tag{11}$$

Finally, we define the signed Sentiment Utopia Index corresponding to the particular choice of weight function as

$$\mathrm{SUI}^{t,\mathsf{E}}(w) := \mathrm{SUI}^{(+)}_{t,\mathsf{E}}(w) - \mathrm{SUI}^{(-)}_{t,\mathsf{E}}(w) \in [-1,1], \tag{12}$$

where positive values indicate bullish orientation and negative values indicate bearish orientation, each measured relative to the historically feasible envelope bounds.

2.8 Meta-ordering interpretation of sentiment profiles

The construction of the sentiment envelopes admits a natural interpretation in terms of dominance relations on the space of sentiment profiles. This interpretation sheds some order theoretic light in both the role of the envelopes and the meaning of the Sentiment Utopia Index as a quantitative measure of distance from dominance.

Let $\sigma_k(\cdot)$ and $\sigma_{k'}(\cdot)$ be two sentiment profiles associated with documents d_k and $d_{k'}$, respectively. We define a pointwise dominance relation by

$$d_k \succcurlyeq d_{k'} \quad \Longleftrightarrow \quad \sigma_k(x) \geq \sigma_{k'}(x) \quad \text{for all } x \in [0,1].$$

This defines a preorder on the space of sentiment profiles: a profile dominates another if it is everywhere at least as bullish (or at least as positive in tone) at every rhetorical position. This order is generally incomplete, since most pairs of profiles will cross and therefore be incomparable.

The upper envelope $\overline{\sigma}(x)$ and lower envelope $\underline{\sigma}(x)$ defined in (7) and (10) play the role of extremal reference profiles in this pre-order. The upper envelope corresponds to a corpus-wide maximally optimistic sentiment profile: a hypothetical profile that is at least as bullish as every observed document at every rhetorical position, and therefore dominates every profile in the corpus. Analogously, the lower envelope corresponds to a dystopian profile that is everywhere as bearish as possible relative to the corpus and is dominated by every profile. Neither envelope need correspond to any actual document; they are order idealized objects induced by the partial order itself.

The envelopes therefore define bounds on the feasible sentiment region of the corpus, and they induce a meta-ordering problem: while direct dominance is generically rare, one may ask how far a given profile is from dominating the corpus, or from being dominated by it. The Sentiment Utopia Index answers precisely this question: Given weight function $w(\cdot)$, the positive shortfall functional $\mathrm{SF}^+_k(w)$ represents the weighted obstacle to d_k dominating the corpus: it aggregates the extent to which σ_k falls short of the utopian envelope at each rhetorical position. Its normalization by the maximal feasible shortfall yields $\mathrm{SUI}^{(+)}_k(w)$ which lies in $[0,1]$ and equals one if and only if $\sigma_k(x) = \overline{\sigma}(x)$ almost everywhere. Thus $\mathrm{SUI}^{(+)}$ can be interpreted as a degree of dominance proximity; it provides a quantification of how un-obstructed is a document to being corpus-dominant in the sentiment pre-order.

The interpretation of $\mathrm{SUI}_k^{(-)}(w)$ is dual; it represents proximity to the sentimentally dystopian extreme. The signed index $\mathrm{SUI}_k(w)$ therefore locates each document within the corpus-relative sentiment spectrum, with positive values indicating sentimental tilting towards the bullish extreme and negative values towards the bearish extreme respectively.

This construction is formally analogous to dominance-based indices used in other domains, such as stochastic dominance in income distributions (Hadar and Russell, 1969; Anderson et al., 2020), and to the curvature-based dominance and utopia constructions used in our earlier work on structural cohesion (Arvanitis, 2025). In all cases, envelopes represent extremal elements induced by a partial order, and utopia indices quantify obstacles from those extremal elements in a way that preserves interpretability, comparability, and robustness to scale.

Under this interpretation, the Sentiment Utopia Index is not merely a normalized score but a functional meta-ordering device: it embeds a set of high-dimensional sentiment profiles into a one-dimensional scale that reflects their position relative to corpus-wide rhetorical extremes, while preserving the underlying dominance logic that gives the index its meaning.

Obviously this induced meta-ordering is not unique, but depends on the choice of the weighting function $w(\cdot)$. Different weights correspond to different implicit notions of rhetorical importance across the document and therefore generate different quantitative refinements of the same underlying dominance preorder. In this sense, $w(\cdot)$ plays the role of a preference or attention functional over rhetorical positions: it determines which obstacles to dominance are emphasized and which are downweighted.

This dependence admits both agnostic and data-driven interpretations. A uniform weight corresponds to a maximum-entropy or least-informative choice, treating all rhetorical positions symmetrically and yielding a baseline, preference-free refinement of the dominance structure. Alternatively, weights may be inferred from data using machine learning or econometric criteria, for example by selecting $w(\cdot)$ to optimize predictive alignment with an economic outcome. Such approaches can be viewed as learning a task-specific refinement of the sentiment meta-ordering, while preserving the same envelope-induced dominance logic that underlies the construction of the Sentiment Utopia Index.

2.9 Time scales and aggregation structure.

The construction of the Sentiment Utopia Index involves three distinct dimensions that should be kept conceptually separate. First, there is *calendar time* t, which indexes the actual arrival of documents and determines the temporal ordering relevant for forecasting and real-time information sets. Second, there is *rhetorical time* $x \in [0, 1]$, which indexes the normalized position of sentences within a document and captures the internal rhetorical progression of sentiment. Third, there is a *cross-sectional* dimension arising from the presence of multiple documents—potentially from different firms—released on the same calendar date.

Aggregation proceeds sequentially across these dimensions. Sentence-level sentiment scores are first aggregated within each document across rhetorical time to form a functional sentiment profile $\sigma_{t,j}(\cdot)$. When multiple documents arrive on the same day t, their profiles are aggregated across the cross section to produce a single daily profile $\sigma_t(\cdot)$. Corpus-level envelopes are then constructed across historical calendar time, yielding upper and lower reference profiles that define the feasible sentiment region at each rhetorical position. Finally, the Sentiment Utopia Index collapses the functional object $\sigma_t(\cdot)$ into a scalar by integrating weighted deviations from these envelopes over rhetorical time. This layered aggregation ensures that rhetorical structure, cross-sectional variation, and temporal ordering are all explicitly accounted for, while preserving a real-time interpretation suitable for forecasting applications.

2.10 Applications, scope, and extensions

The above yield a functional sentiment $\mathcal{T}$-time series $\{\sigma_t(\cdot)\}$ together with a family of scalar indices $\left\{\mathrm{SUI}^{t,\mathsf{E}}(w_\theta)\right\}$ that summarize corpus-relative extremeness under alternative rhetorical emphasis patterns and envelope dynamics. These objects can be used in several empirical roles. First, they support descriptive monitoring of narrative regimes through the level, volatility, and persistence of $\mathrm{SUI}^{t,\mathsf{E}}(w_\theta)$ and through direct inspection of the underlying profiles $\sigma_t(\cdot)$. Second, they support structural analysis of rhetorical sentiment evolution by treating $\sigma_t(\cdot)$ as a functional time series-see for example Kokoszka and Reimherr (2017)-and studying how the shape of sentiment trajectories changes over time. Third, they support forecasting applications by treating the SUI as an information variable in predictive models. The empirical application in §3 focuses on this third role and evaluates whether SUI provides incremental predictive content for realized volatility beyond a standard HAR benchmark; the forecasting specification, windowing protocol, loss functions, and forecast comparison tests are therefore presented in the empirical section so that the present methodology section remains centered on the construction of the sentiment objects themselves.

The framework is modular and admits a range of natural extensions that could be conceptually interesting as objects of further theoretical and applied research.

A first class of extensions concerns the treatment of rhetorical weighting. In the baseline design, rhetorical emphasis is treated as a robustness dimension: we compute a family of indices over a fixed parametric grid rather than estimating weights from labeled data or optimizing weights for a specific economic target. A natural extension would be to shift from sensitivity analysis to estimation. One possibility is to consider nonparametric classes of admissible weights (e.g., classes of monotone or shape-restricted weights) that capture early- versus late-emphasis reading schemes without imposing a specific parametric form. Another is to estimate weights by optimizing predictive alignment with a target variable (e.g., realized volatility, returns, or option-implied measures), yielding economically trained rhetorical weights. A further extension would allow weights to evolve over time, reflecting changes in market attention, disclosure practices, or narrative style across regimes. Such approaches would move rhetorical emphasis from an uncertainty dimension into an adaptive or structural component of the model, but they also introduce additional estimation uncertainty, model selection issues, and heightened overfitting risk; consequently, they require explicit training/validation separation and, ideally, nested cross-validation.

A second set of extensions concerns the construction and interpretation of the normalization envelopes. In the baseline design, envelopes are computed either in an expanding or fixed-length rolling manner and serve solely to define the historical bounds relative to which extremeness is measured. However, envelope dynamics may themselves carry economic information: the envelope width provides a corpus-relative measure of how dispersed narrative sentiment can be at each rhetorical location, and its evolution can reflect shifts in the range of rhetorical expression. Allowing envelope width summaries (or their changes) to enter predictive regressions would permit the analysis of narrative dispersion as an additional channel beyond extremeness. Similarly, varying the rolling window length would allow the normalization frame to adapt at different temporal scales, capturing short-lived narrative regimes or long-run structural shifts. Finally, change-point detection applied to envelope dynamics offers a principled way to identify and temporally localize narrative regime changes and to study whether the predictive role of sentiment extremeness is regime dependent.

A further extension involves the functional structure of sentiment itself. The current approach compresses each day's information into a single profile $\sigma_t(\cdot)$ and then further compresses it into scalar indices. This compression is attractive for forecasting, but it could discard informative functional variation. Extensions could incorporate functional principal components of $\sigma_t(\cdot)$, dynamic functional factor models in which latent narrative dimensions evolve over time, or cross-sectional dispersion measures (e.g., across firms or sectors) to capture disagreement and heterogeneity in narratives. These directions would shift the analysis from scalar indices toward genuinely high-dimensional functional state variables, potentially enabling richer structural inference at the cost of additional modeling complexity.

Finally, another natural direction concerns the scope and granularity of the text universe. The current implementation uses only periodic SEC filings and aggregates sentiment at the market level. Future work could expand the corpus to include larger universes of documents like 8-K filings, earnings-call transcripts, press releases, news articles, or tweets, each of which reflects distinct informational channels and temporal dynamics. Multilingual corpora would enable the study of international narrative transmission and spillovers. Firm-level or sector-level implementations would permit cross-sectional analysis of narrative effects, while topic-conditioned indices could isolate specific themes which could for example be informational to empirical questions like time-stamping bubble formation and bursting beyond price dynamics.

3 Empirical Application

This section provides an empirical investigation of whether the SUI, derived from rhetorical sentiment structures in financial narratives, contains predictive information about (near) future financial market volatility. Specifically, we ask whether SUI variants, constructed with different parametric rhetorical weightings and evaluated under envelope-normalized sentiment profiles, improve out-of-sample forecasts of realized volatility beyond standard benchmarks.

We implement a real-time forecasting exercise in which each SUI series is incorporated into a HAR-type model for realized volatility–see Corsi (2009). Forecast performance is evaluated using both rolling and expanding estimation windows, two volatility loss functions (MSE and QLIKE), and formal loss-differential tests (Diebold–Mariano and Clark–West–see Diebold and Mariano (1995) and Clark and West (2007) respectively). Our objective is not necessarily to identify a single optimal SUI configuration, but to assess whether rhetorically-structured narrative extremeness yields robust signals of future volatility that complement lagged volatility dynamics.

The empirical analysis proceeds as follows. Section 3.1 describes the implementation specifics of the SUI methodology: (a.) the construction of the SEC filing corpus and the alignment of texts to trading days, (b.) the natural language processing pipeline, including sentence segmentation and sentiment scoring, (c.) the construction of document-and day-level sentiment profiles via discretization, (d.) the envelope-based normalization of bullish and bearish sentiment and the evaluation of the SUI across a parametric family of rhetorical weights. Section 3.2 describes the forecasting model specification, the real-time protocol, and the statistical evaluation criteria. Section 3.3 briefly discusses computational issues. Finally, Section 3.4 presents the results, including forecast performance rankings of SUI variants, significance testing, and some interpretive discussion.

It is important to note that throughout, we treat the rhetorical weighting specification as a structured form of model uncertainty and assess the robustness of forecasting gains across alternative envelope schemes and rhetorical emphasis profiles.

3.1 SUI implementation specifics

The empirical corpus consists of SEC EDGAR filings restricted to 10-K and 10-Q forms; we do not use 8-K filings in the present implementation). This restriction is intentional: 8-Ks include heterogeneous event-driven disclosures and can introduce sharp timing-induced selection effects in sentiment profiles, while the 10-K/10-Q stream is more regular, comparable, and better aligned to slow-moving narrative structure.

Each filing is mapped to a trading-day index using the chosen market calendar (NYSE). When multiple filings fall on the same trading day, the baseline aggregation rule is the average-profile rule (daily profile equals the average of within-day document profiles), consistent with the paper's baseline aggregation proposal. Each document is parsed into sentences using `spaCy`. Sentences form the primitive units for constructing the document sentiment trajectory $\sigma_k(\cdot)$ on $[0, 1]$. To stabilize runtime and memory consumption (and to avoid pathological parsing of extremely long lines in filings), we impose two explicit caps:

- Maximum sentences per document: 400 (documents longer than 400 sentences are truncated).
- Maximum characters per sentence: 400 (sentences longer than 400 characters are truncated).

The caps are useful when scaling to multiple firms or longer sample ranges.

Sentence-level sentiment is computed using FinBERT outputs as $\sigma_i = p_{pos} - p_{neg}$, producing a signed sentiment score per sentence. This is then lifted to the rhetorical-time step function $\sigma_k(x)$ by assigning each sentence to its corresponding subinterval of $[0, 1]$.

Each document profile $\sigma_k(\cdot)$ is evaluated on a uniform grid of size $G = 100$, yielding a vector representation $\{\sigma_k(x_g)\}_{g=1}^{G}$ for envelope construction and SUI evaluation. Also, at each t index trading days and let $\mathcal{D}_t$ denote the set of filings aligned to day t. The daily profile is computed as the average of document profiles in $\mathcal{D}_t$.

The SUI is constructed using corpus-level bullish and bearish envelopes. In forecasting contexts, envelopes must respect real-time information constraints, motivating the expanding and rolling schemes.[1] Also, as noted in the previous section, we compute SUI series under

[1] The envelope definitions and their forecasting interpretation are discussed formally in the paper; see especially the expanding and rolling constructions and their comparability tradeoffs.

expanding and rolling envelopes; the latter at each time t uses use only the trailing $K = 225$ trading days, adapting to local regimes.

We evaluate a structured grid of parametric weight functions $w_\theta(x) = \exp(\theta_1 x + \theta_2 x^2)$, indexed by the parameter vector $\theta = (\theta_1, \theta_2) \in \Theta_G \subset \mathbb{R}^2$, where Θ_G denotes the finite evaluation grid used for empirical implementation. Each parameter combination $\theta \in \Theta_G$ defines a distinct rhetorical weighting profile over normalized rhetorical time $x \in [0, 1]$, which is then used to compute the signed SUI series under each envelope scheme $E \in \{\exp, \mathrm{roll}\}$. This yields a panel of sentiment indices $\{\mathrm{SUI}_{t,E}(w_\theta)\}_{\theta \in \Theta_G}$ for each weighting configuration and envelope method. In our implementation, we use the grid

$$\Theta_G = \{(\theta_1, \theta_2) \in \{-2, -1, 0, 1, 2\} \times \{-2, 0, 2\}\},$$

resulting in $5 \times 3 = 15$ distinct rhetorical weighting functions. For each $\theta \in \Theta_G$, the weights $w_\theta(x)$ are discretized over a fixed grid of rhetorical time points $\{x_g\}_{g=1}^{G}$, normalized to sum to one, and then used to compute time-varying weighted sentiment profiles. The resulting SUI columns are matched across envelope schemes to ensure comparability in the subsequent forecasting evaluation.

3.2 Volatility proxy and forecast design

This subsection describes the construction of the volatility target, the forecasting models, the out-of-sample protocol, and the statistical evaluation procedures used in the empirical analysis.

Daily volatility is proxied using an OHLC-based estimator-see for example Garman and Klass (1980); Yang and Zhang (2000)-constructed from daily open, high, low, and close prices obtained from Stooq for the S&P 500 index. Let O_t, H_t, L_t, and C_t denote the open, high, low, and close prices on day t. The baseline estimator used in the implementation is the Garman–Klass estimator-see Garman and Klass (1980),

$$\mathrm{RV}_t^{\mathrm{GK}} = \tfrac{1}{2}\big(\log(H_t/L_t)\big)^2 - (2\log 2 - 1)\big(\log(C_t/O_t)\big)^2,$$

which is consistent and efficient under the assumption of continuous sample paths with no overnight jumps and is standard in the empirical volatility literature. Alternative estimators (e.g. Parkinson or close-to-close–see for example Andersen et al. (2003)) can be readily incorporated in the pipeline.

Given the strict positivity of the volatility proxy, in the forecasting regressions we use the logarithmic transformation

$$Y_{t+1} = \log(\mathrm{RV}_{t+1}),$$

which stabilizes variance, reduces right-skewness, and aligns with the approximately lognormal empirical distribution of realized volatility. All forecasts are therefore produced on the log-volatility scale and then mapped back to levels when QLIKE losses are computed.

Benchmark and augmented forecasting models. The benchmark forecasting model is the heterogeneous autoregressive (HAR) model of realized volatility, which captures the long-memory structure of volatility through daily, weekly, and monthly components. Define

$$RV_t^{(d)} = RV_t, \qquad RV_t^{(w)} = \frac{1}{5}\sum_{j=0}^{4} RV_{t-j}, \qquad RV_t^{(m)} = \frac{1}{22}\sum_{j=0}^{21} RV_{t-j}.$$

The benchmark model is

$$\log RV_{t+1} = \alpha + \beta_1 \log RV_t^{(d)} + \beta_2 \log RV_t^{(w)} + \beta_3 \log RV_t^{(m)} + u_{t+1}.$$

The augmented model adds a single sentiment regressor:

$$\log RV_{t+1} = \alpha + \beta_1 \log RV_t^{(d)} + \beta_2 \log RV_t^{(w)} + \beta_3 \log RV_t^{(m)} + \delta\, \mathrm{SUI}_t(w) + u_{t+1}.$$

This specification isolates the incremental predictive content of the sentiment index relative to the strong autoregressive baseline. Only one SUI regressor is included at a time in order to avoid multicollinearity across different rhetorical weightings and to ensure interpretability of δ as the marginal effect of sentiment extremeness.

Out-of-sample forecasting protocol. Forecasts are generated recursively in a pseudo-real-time manner. Let t_0 denote the first forecast origin. For each $t = t_0, \ldots, T-1$:

1. The model parameters are estimated using only observations available up to time t, using either an expanding or rolling estimation window.
2. A one-step-ahead forecast $\widehat{Y}_{t+1}$ is produced.
3. The realized Y_{t+1} is observed and forecast errors are recorded.

In the baseline implementation we use:

- a rolling estimation window of fixed length $R = 500$ trading days,
- first forecast origin $t_0 = 500$,
- identical windowing and data availability constraints for both benchmark and augmented models.

This design ensures that the augmented model is evaluated under exactly the same informational and sample conditions as the benchmark, so that differences in performance can be attributed solely to the inclusion of the SUI regressor.

Loss functions and forecast evaluation. Forecast performance is evaluated using two complementary loss functions.

On the log scale we report mean squared error,

$$L_{\mathrm{MSE}}(Y_{t+1}, \widehat{Y}_{t+1}) = (Y_{t+1} - \widehat{Y}_{t+1})^2,$$

which penalizes large forecast errors symmetrically.

On the level scale we report the QLIKE loss (see for example Patton (2011)),

$$L_{\mathrm{QLIKE}}(RV_{t+1}, \widehat{RV}_{t+1}) = \log(\widehat{RV}_{t+1}) + \frac{RV_{t+1}}{\widehat{RV}_{t+1}},$$

which is robust to measurement error in volatility proxies and is widely used for evaluating volatility forecasts.

Statistical tests for predictive ability. Let $\ell_{t+1} = L_{t+1}^{\text{aug}} - L_{t+1}^{\text{base}}$ denote the loss differential between the augmented and benchmark models. We test the null hypothesis $\mathbb{E}[\ell_{t+1}] = 0$ using the Diebold–Mariano test with heteroskedasticity-and-autocorrelation consistent (HAC) standard errors–see Diebold and Mariano (1995). A negative mean loss differential indicates that the augmented model outperforms the benchmark. Since the augmented model nests the benchmark, we also report the Clark–West adjusted test–see Clark and West (2007), which adjusts the Diebold-Mariano statistic for the upward bias in nested-model forecast comparisons. Both tests are implemented using a HAC bandwidth of $L = 5$ lags.

Under this design, evidence of predictive value is interpreted conservatively. A statistically significant improvement in MSE or QLIKE indicates that rhetorically structured, corpus-normalized sentiment extremeness contains incremental information about future volatility beyond what is captured by volatility persistence alone. Obviously, the design does not claim structural identification of sentiment shocks, nor does it assert that narratives cause volatility in a causal sense; rather, it establishes whether narrative extremeness is an economically meaningful forecasting signal.

3.3 Computational implementation and runtime characteristics

The full empirical pipeline is implemented in Python and is designed to be modular, restartable, and robust to large-scale re-runs and parameter sweeps. All intermediate outputs are cached to disk using columnar storage formats (Parquet for large arrays and CSV for summary tables), so that expensive stages are executed only once and downstream analyses can be repeated without recomputation. In the current baseline run covering the 2021–2024 sample window for a single aggregate market series, the dominant computational cost arises from transformer-based sentence-level sentiment scoring, which operates at the level of individual sentences and therefore scales with both document length and corpus size. This stage typically requires several hours on a standard CPU machine and is the primary bottleneck in the pipeline. In contrast, earlier stages such as network retrieval of filings are largely I/O-bound and incur negligible cost after the first cached run, while subsequent stages such as sentence aggregation into rhetorical profiles, construction of daily profiles, and envelope computation operate on fixed-size arrays and complete in minutes or less. The generation of multiple SUI variants over the parametric weight grid and envelope schemes is computationally lightweight and takes seconds to minutes, and the out-of-sample forecasting and loss evaluation loops scale approximately linearly with the number of SUI variants and complete in minutes for the full grid.

Notice that the two truncation rules imposed at the pre-processing stage–see Section 3.1–substantially reduce the heavy-tailed distribution of document lengths and make the pipeline stable enough to support large re-runs, rolling-window recomputation, and systematic robustness analysis across rhetorical weightings and envelope specifications.

3.4 Results

We rank each candidate SUI series $\text{SUI}_{t,E}(w_\theta)$ by its incremental out-of-sample forecast performance in the HAR–SUI model relative to the HAR benchmark. Because MSE and

QLIKE emphasize different aspects of forecast accuracy, we report two rankings: (i) by augmented-model OOS MSE, and (ii) by augmented-model OOS QLIKE.[2]

Across all configurations, the HAR benchmark delivers an OOS MSE of approximately 0.6594 in $Y = \log(\mathrm{RV})$ units. The best SUI-augmented specifications reduce MSE by about 0.004–0.006 (roughly <1% improvement), and improve QLIKE by about 0.006–0.007 in absolute units. The gains are economically small in magnitude but notably structured across the weight grid.

Tables 1a–2a report the top candidates under each envelope scheme. Cells in green indicate improvements (negative Δ loss); red indicates deterioration.

(a) Top 10 by augmented OOS MSE

Table (1) Top SUI variants under **expanding** envelopes (K not applicable), ranked by MSE and by QLIKE.

SUI	θ_1	θ_2	MSE_1	ΔMSE	QLIKE_1	ΔQLIKE	DM p	CW p
sui_t1-2p0_t2-2p0	-2.0	-2.0	0.6535	-0.0059	−9.1840	-0.0066	0.112	0.033
sui_t1-1p0_t2-2p0	-1.0	-2.0	0.6539	-0.0055	−9.1850	-0.0069	0.130	0.043
sui_t1+0p0_t2-2p0	0.0	-2.0	0.6550	-0.0044	−9.1840	-0.0066	0.178	0.096
sui_t1-2p0_t2+0p0	-2.0	0.0	0.6551	-0.0043	−9.1840	-0.0066	0.192	0.102
sui_t1+1p0_t2-2p0	1.0	-2.0	0.6570	-0.0025	−9.1830	-0.0053	0.304	0.219
sui_t1-1p0_t2+0p0	-1.0	0.0	0.6572	-0.0022	−9.1830	-0.0051	0.323	0.230
sui_t1+2p0_t2+2p0	2.0	2.0	0.6579	-0.0015	−9.1760	0.0014	0.358	0.287
sui_t1+1p0_t2+2p0	1.0	2.0	0.6584	-0.0010	−9.1760	0.0018	0.409	0.341
sui_t1+2p0_t2-2p0	2.0	-2.0	0.6592	-0.0002	−9.1800	-0.0027	0.488	0.430
sui_t1+0p0_t2+2p0	0.0	2.0	0.6592	-0.0002	−9.1760	0.0020	0.496	0.438

(b) Top 10 by augmented OOS QLIKE

SUI	θ_1	θ_2	QLIKE_1	ΔQLIKE	MSE_1	ΔMSE	DM p	CW p
sui_t1-1p0_t2-2p0	-1.0	-2.0	−9.1850	-0.0069	0.6539	-0.0055	0.130	0.043
sui_t1-2p0_t2-2p0	-2.0	-2.0	−9.1840	-0.0066	0.6535	-0.0059	0.112	0.033
sui_t1+0p0_t2-2p0	0.0	-2.0	−9.1840	-0.0066	0.6550	-0.0044	0.178	0.096
sui_t1-2p0_t2+0p0	-2.0	0.0	−9.1840	-0.0066	0.6551	-0.0043	0.192	0.102
sui_t1+1p0_t2-2p0	1.0	-2.0	−9.1830	-0.0053	0.6570	-0.0025	0.304	0.219
sui_t1-1p0_t2+0p0	-1.0	0.0	−9.1830	-0.0051	0.6572	-0.0022	0.323	0.230
sui_t1+2p0_t2-2p0	2.0	-2.0	−9.1800	-0.0027	0.6592	-0.0002	0.488	0.430
sui_t1+0p0_t2+0p0	0.0	0.0	−9.1800	-0.0021	0.6596	0.0001	0.525	0.468
sui_t1-2p0_t2+2p0	-2.0	2.0	−9.1790	-0.0013	0.6599	0.0005	0.551	0.496
sui_t1+1p0_t2+0p0	1.0	0.0	−9.1770	0.0008	0.6604	0.0009	0.606	0.559

[2]QLIKE is particularly standard for volatility forecasting because it penalizes scale errors asymmetrically and is robust to heavy-tailed realized-volatility distributions.

(a) Top 10 by augmented OOS MSE

Table (2) Top SUI variants under **rolling** envelopes ($K = 252$), ranked by MSE and by QLIKE.

SUI	θ_1	θ_2	MSE_1	ΔMSE	QLIKE_1	ΔQLIKE	DM p	CW p
sui_t1-2p0_t2-2p0	-2.0	-2.0	0.6537	-0.0057	−9.1840	-0.0065	0.128	0.040
sui_t1-1p0_t2-2p0	-1.0	-2.0	0.6539	-0.0055	−9.1840	-0.0068	0.137	0.046
sui_t1+0p0_t2-2p0	0.0	-2.0	0.6551	-0.0043	−9.1840	-0.0065	0.185	0.104
sui_t1-2p0_t2+0p0	-2.0	0.0	0.6551	-0.0043	−9.1840	-0.0065	0.189	0.106
sui_t1+1p0_t2-2p0	1.0	-2.0	0.6571	-0.0023	−9.1830	-0.0052	0.309	0.227
sui_t1-1p0_t2+0p0	-1.0	0.0	0.6573	-0.0021	−9.1830	-0.0051	0.329	0.242
sui_t1+2p0_t2+2p0	2.0	2.0	0.6580	-0.0014	−9.1760	0.0015	0.366	0.296
sui_t1+1p0_t2+2p0	1.0	2.0	0.6585	-0.0009	−9.1760	0.0019	0.415	0.349
sui_t1+2p0_t2-2p0	2.0	-2.0	0.6593	-0.0001	−9.1800	-0.0026	0.492	0.436
sui_t1+0p0_t2+2p0	0.0	2.0	0.6593	-0.0001	−9.1760	0.0021	0.500	0.444

(b) Top 10 by augmented OOS QLIKE

SUI	θ_1	θ_2	QLIKE_1	ΔQLIKE	MSE_1	ΔMSE	DM p	CW p
sui_t1-1p0_t2-2p0	-1.0	-2.0	−9.1840	-0.0068	0.6539	-0.0055	0.137	0.046
sui_t1-2p0_t2-2p0	-2.0	-2.0	−9.1840	-0.0065	0.6537	-0.0057	0.128	0.040
sui_t1+0p0_t2-2p0	0.0	-2.0	−9.1840	-0.0065	0.6551	-0.0043	0.185	0.104
sui_t1-2p0_t2+0p0	-2.0	0.0	−9.1840	-0.0065	0.6551	-0.0043	0.189	0.106
sui_t1+1p0_t2-2p0	1.0	-2.0	−9.1830	-0.0052	0.6571	-0.0023	0.309	0.227
sui_t1-1p0_t2+0p0	-1.0	0.0	−9.1830	-0.0051	0.6573	-0.0021	0.329	0.242
sui_t1+2p0_t2-2p0	2.0	-2.0	−9.1800	-0.0026	0.6593	-0.0001	0.492	0.436
sui_t1+0p0_t2+0p0	0.0	0.0	−9.1800	-0.0020	0.6596	0.0002	0.530	0.476
sui_t1-2p0_t2+2p0	-2.0	2.0	−9.1790	-0.0012	0.6600	0.0006	0.556	0.503
sui_t1+1p0_t2+0p0	1.0	0.0	−9.1770	0.0009	0.6605	0.0010	0.611	0.566

The results may point to pattern across both expanding and rolling envelope schemes. The leading configurations seem to consistently concentrate around $\theta_2 \approx -2$, with $\theta_1 \in \{-2, -1, 0\}$. In this region the weighting function $w_\theta(x)$ is log-concave and monotonically decreasing on $[0, 1]$, with its maximum attained at the beginning of the document. Consequently, predictive content is strongest when rhetorical weighting places disproportionate emphasis on early sentiment realizations.

Economically, this may suggest that volatility reacts primarily to sentiment extremeness that appears early in disclosures, consistent with framing and anchoring effects whereby initial rhetorical tone shapes market interpretation of subsequent information. Mild or late-stage sentiment fluctuations carry substantially less predictive content once early tone is fixed.

Second, the top sets under expanding and rolling envelopes are nearly identical. This matters because expanding envelopes deliver strict real-time comparability, while rolling envelopes adapt to local regimes at the cost of changing normalization. The fact that the same region of (θ_1, θ_2) performs well under both may suggest that the signal is not an artifact of a particular normalization regime.

The Clark–West one-sided p-values for the best specifications are often meaningfully smaller than the corresponding DM p-values. This is expected: HAR+SUI nests HAR, and the CW adjustment corrects for the upward bias in the larger model's MSFE under the null of no incremental predictive content. The CW test therefore provides the more appropriate significance lens.

4 Conclusions

We propose a functional, corpus-relative sentiment measure that (i) reflects rhetorical structure by representing each filing as a sentiment function over normalized rhetorical time, and (ii) enforces historical and contemporaneous normalization through data-driven bullish and bearish envelopes computed from the evolving corpus. This construction yields sentiment characterizations whose interpretation is inherently relative to what has been historically (or temporally local) "possible" in the narrative record, rather than to an arbitrary scale of raw classifier scores. The resulting Sentiment Utopia Index (SUI) provides a way to track when a time instance's narrative profile is unusually optimistic or pessimistic given the history of the corpus, while remaining compatible with real-time forecasting constraints via expanding or rolling envelope schemes.

A further feature of the particular approach is that we treat rhetorical emphasis as a structured form of model uncertainty. Instead of committing to a single weighting of rhetorical positions, we evaluate a parametric grid of weight functions and treat the implied SUI series as a family of candidate narrative signals. In the empirical application, this uncertainty-aware perspective is vindicated by the stability of the ranking patterns across (i) alternative envelope schemes and (ii) alternative forecast loss functions (MSE and QLIKE). The leading configurations cluster in the same region of the weight-parameter space—favoring strong emphasis on early deviations from the sentiment envelope, with rhetorical weight concentrated at the beginning of disclosures—suggesting that the incremental signal is not an artifact of a particular normalization regime or loss criterion. Although the magnitude of out-of-sample gains is modest in this indicative application, the consistency of the patterns across evaluation lenses supports the interpretation that the SUI could be capturing a real, economically interpretable narrative component linked to forecastability of volatility dynamics.

More broadly, the functional time-series framing enables at least two complementary uses of financial narratives. First, it supports predictive analysis, where SUI enters standard forecasting designs (e.g., HAR-type volatility regressions) and can be evaluated using econometric tools for nested forecast comparison. Second, it may also support structural analysis of narratives, since the same objects that forecast can also be studied as evolving functional profiles: one can characterize where in rhetorical time optimism/pessimism concentrates, how envelope bounds drift across regimes, and how narrative "shape" changes around macro or firm-specific events.

Several extensions are immediate. First, rhetorical weights can be learned rather than gridded: one can estimate parametric or nonparametric $w(\cdot)$ using human coherence labels, analyst annotations, or forecast-optimality criteria under cross-validation, while preserving the envelope-based normalization. Second, multilingual corpora and cross-jurisdiction disclosure regimes can be incorporated, enabling comparative narrative dynamics across markets and languages. Finally, richer functional dynamics can be modeled directly—e.g., functional autoregressions, regime-switching envelopes, or state-space representations—so that narrative evolution is treated as a first-class object rather than a reduced-form covariate.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Anderson, G., Post, T., and Whang, Y.-J. (2020). Somewhere between utopia and dystopia: Choosing from multiple incomparable prospects. *Journal of Business & Economic Statistics*, 38(3):502–515.

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models.

Arvanitis, S. (2025). Measuring textual cohesion via graph curvature, weighted dominance, and embedding-based coherence.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Bybee, L., Kelly, B. T., Manela, A., and Xiu, D. (2024). Business news and business cycles. *The Journal of Finance*, 79(5):3105–3147.

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.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1):67–78.

Hadar, J. and Russell, W. R. (1969). Rules for ordering uncertain prospects. *The American economic review*, 59(1):25–34.

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

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., and Kapadia, S. (2022). Making text count: economic forecasting using newspaper text. *Journal of Applied Econometrics*, 37(5):896–919.

Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. Texts in Statistical Science. Chapman and Hall/CRC.

Liu, Z., Huang, D., Huang, K., Li, Z., and Zhao, J. (2021). Finbert: a pre-trained financial language representation model for financial text mining. In *Proceedings of the 29th International Conference on International Joint Conferences on Artificial Intelligence (IJCAI '20)*, pages –. International Joint Conferences on Artificial Intelligence.

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

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2019). Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168.

Yang, D. and Zhang, Q. (2000). Drift-independent volatility estimation based on OHLC prices. *Journal of Business*, 73(3):477–491.

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2024

01-24 **Market Timing & Predictive Complexity, Stelios Arvanitis, Foteini Kyriazi, Dimitrios Thomakos**
02-24 **Multi-Objective Frequentistic Model Averaging with an Application to Economic Growth, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, Nikolas Topaloglou**
03-24 **State dependent fiscal multipliers in a Small Open Economy, Xiaoshan Chen, Jilei Huang, Petros Varthalitis**
04-24 **Public debt consolidation: Aggregate and distributional implications in a small open economy of the Euro Area, Eleftherios-Theodoros Roumpanis**
05-24 **Intangible investment during the Global Financial Crisis in the EU, Vassiliki Dimakopoulou, Stelios Sakkas and Petros Varthalitis**
06-24 **Time will tell! Towards the construction of instantaneous indicators of different agent-types, Iordanis Kalaitzoglou, Stelios Arvanitis**
07-24 **Norm Constrained Empirical Portfolio Optimization with Stochastic Dominance: Robust Optimization Non-Asymptotics, Stelios Arvanitis**
08-24 **Asymptotics of a QLR-type test for optimal predictive ability, Stelios Arvanitis**
09-24 **Strongly Equitable General Equilibrium Allocations, Angelos Angelopoulos**
10-24 **The Greek macroeconomy: A note on the current situation and future outlook, Apostolis Philippopoulos**
11-24 **Evolution of Greek Tax System, A Survey of Legislated Tax Changes from 1974 to 2018, Panagiotis Asimakopoulos**
12-24 **Macroeconomic Impact of Tax Changes, The case of Greece from 1974 to 2018, Panagiotis Asimakopoulos**
13-24 **`Pareto, Edgeworth, Walras, Shapley' Equivalence in a Small Economy, Angelos Angelopoulos**
14-24 **Stimulating long-term growth and welfare in the U.S, James Malley and Apostolis Philippopoulos**
15-24 **Distributionally Conservative Stochastic Dominance via Subsampling, Stelios Arvanitis**
16-24 **Before and After the Political Transition of 1974. Institutions, Politics, and the Economy of Post-War Greece, George Alogoskoufis**
17-24 **Gaussian Stochastic Volatility, Misspecified Volatility Filters and Indirect Inference Estimation, Stelios Arvanitis, Antonis Demos**
18-24 **Endogenous Realtor Intermediation and Housing Market Liquidity, Miroslav Gabrovski, Ioannis Kospentaris, Victor Ortego-Marti**
19-24 **Universal Choice Spaces and Expected Utility: A Banach-type Functorial Fixed Point, Stelios Arvanitis, Pantelis Argyropoulos, Spyros Vassilakis**

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5