Data-Revision Wedges in Real-Time Uncertainty Indices

January 22, 2026

Abstract

We study real-time measurement of macroeconomic uncertainty when the underlying data are revised. In a vintage-data factor forecasting environment, a conventional real-time y-based uncertainty index admits a data-revision wedge, so y-based uncertainty computed from vintages can be mechanically inflated relative to its final-data benchmark. Motivated by this wedge, we propose a real-time factor-based index that targets the factor-innovation component directly and show that its benchmark gap is free of revision-driven terms. Equivalently, the real-time versus final-data gap of the proposed index reflects only information-set differences in factor-innovation uncertainty, whereas the corresponding gap of the y-based index includes the data-revision wedge. The results deliver simple diagnostics that quantify benchmark recovery and revision-induced inflation in real-time uncertainty indices.

1 Introduction

Macroeconomic uncertainty is widely viewed as an important driver of economic activity and a key state variable for policy analysis. When uncertainty rises, firms delay investment and hiring, households postpone durable consumption, financial conditions tighten, and policymakers confront a more fragile environment in which standard signals become noisier. These considerations have motivated a large empirical literature that seeks to measure aggregate uncertainty in a way that is forward-looking, comparable over time, and grounded in observable data.

A prominent approach, introduced by JuradoLudvigsonNg2015, measures macroeconomic uncertainty as the conditional second moment of multi-horizon forecast errors across a large panel of macroeconomic and financial time series. The JLN construction is attractive for two reasons. First, it is explicitly tied to predictability: uncertainty is high when forecast errors are expected to be volatile, given current information. Second, it aggregates information across many series and, in principle, can isolate an economy-wide component rather than idiosyncratic noise in any single variable. In practice, however, JLN implement their index using fully revised historical data rather than real-time vintages.

Our paper takes the practical forecasting perspective as a starting point. In real time, agents do not observe the final revised dataset; they observe a sequence of vintages that are subsequently revised. This raises a basic question for real-time monitoring: *what does a JLN-style forecast-error uncertainty index measure when it is computed using vintage data that will later be revised?* Put differently, if one applies the JLN recipe at time t using the information actually available at time t, does the resulting real-time index track the same primitive object as the final-data benchmark, or can it be mechanically distorted by revisions?

This paper shows that the answer depends critically on factor extraction in real-time forecasting. In large panels, forecasts are often constructed using estimated common factors. When the underlying panel is observed in vintages, data revisions perturb the extracted factor states. These revision-induced perturbations generate *factor-extraction error* that enters implemented forecast errors even when the primitive *factor-innovation uncertainty* is unchanged. As a result, a real-time implementation of a y-based forecast-error uncertainty index can be mechanically inflated by a *revision wedge*—a component of measured uncertainty induced by revisions rather than by genuine aggregate uncertainty.

To make this mechanism transparent, we develop a stylized real-time factor forecasting environment in which each vintage observation admits a decomposition into a common-factor component, an idiosyncratic component, and a revision component. Within this setup, we derive sharp accounting relationships for environment-specific y-based uncertainty. These relationships decompose measured uncertainty into three parts: (i) *factor-innovation uncertainty*, governed by the conditional variance of multi-step factor forecast errors; (ii) an idiosyncratic component; and (iii) a factor-extraction-error component that arises because forecasting uses extracted factor states rather than the latent state. In the real-time environment, the factor-extraction-error component further splits into a non-revision part and a revision-induced part. The latter is the revision wedge and is the unique channel through which data revisions contaminate real-time y-based uncertainty.

The theoretical implications are immediate. First, the real-time versus final-data gap of the conventional y-based index decomposes into an information-set term (reflecting different conditioning sets) plus non-common components that include the revision wedge. Under strong diversification conditions, the revision wedge can attenuate as the cross-sectional dimension grows, but there is no reason for it to be negligible in empirically relevant finite-N settings. Second, and most importantly for practice, the accounting relationships point to an alternative index that targets factor-innovation uncertainty directly and therefore avoids revision-driven inflation.

Motivated by this logic, we propose a *real-time factor-based uncertainty index*, denoted by $U_t^{f,\mathrm{RT}}(h)^2$, defined on the conditional variance of multi-step factor forecast errors and mapped to the scale of the target panel. The proposed index is deliberately positioned relative to the JLN benchmark. In our notation, the benchmark implementation in JuradoLudvigsonNg2015 corresponds to the final-data y-based macro index $U_t^{y,\mathrm{F}}(h)^2$, while the naive real-time counterpart is $U_t^{y,\mathrm{RT}}(h)^2$. Our first key result provides an explicit accounting decomposition for the *benchmark gap* $U_t^{f,\mathrm{RT}}(h)^2 - U_t^{y,\mathrm{F}}(h)^2$ and shows that this gap contains no revision wedge. A complementary result shows that the revision wedge appears in the real-time versus final-data gap of the y-based index but is absent from the corresponding gap of the factor-based index. Together, these results formalize a central message of the paper: $U_t^{f,\mathrm{RT}}(h)^2$ recovers, in real time, the factor-innovation uncertainty component underlying the JLN final-data benchmark, while remaining insulated from revision-induced factor-extraction error.

A practical advantage of the framework is that it yields empirical reporting quantities that map directly to the theory. In the empirical analysis we report (i) benchmark-recovery summaries that quantify how closely $U_t^{f,\mathrm{RT}}(h)^2$ tracks $U_t^{y,\mathrm{F}}(h)^2$; (ii) a difference-in-gaps diagnostic that isolates the non-common real-time component embedded in $U_t^{y,\mathrm{RT}}(h)^2$ relative to the factor-based construction; and (iii) horizon- and cross-sectional-dimension patterns implied by the model (horizon damping and large-N attenuation). These summaries are informative even when the revision wedge itself is not separately identified, and they provide a clear way to assess when revisions matter quantitatively for real-time uncertainty measurement.

Contributions. The paper makes three contributions. First, it provides a tractable real-time factor environment with vintages and revisions that delivers sharp accounting decompositions of forecast-error uncertainty and isolates an explicit revision wedge in real time. Second, it proposes the real-time factor-based index $U_t^{f,\mathrm{RT}}(h)^2$ and establishes benchmark comparisons linking the proposed real-time object to the JLN final-data benchmark, clarifying exactly what is removed by the new construction and what differences remain. Third, it translates the theory into empirical reporting quantities that quantify benchmark recovery and revision-related real-time distortions in y-based uncertainty, and that provide model-implied patterns across horizons and cross-sectional dimension.

Related literature.

Roadmap. Section 2 introduces the real-time factor environment, the vintage structure, and the forecasting setup. Section 3 develops the main accounting identities, derives the real-time versus final-data gap decompositions, and introduces the proposed real-time factor index together with the benchmark comparison results.

2 A Real-Time Factor DGP with Data Revisions

This section formalizes the real-time (vintage-data) forecasting environment used throughout the paper. The key feature is that the predictor panel is observed in vintages and is subject to data revisions, which induces additional noise in real-time factor proxies. Section 3 shows how this revision-driven proxy noise contaminates forecast-error-based uncertainty measures constructed in real time.

2.1 Data environments and information sets

Time is indexed by $t \in \{1, \ldots, T\}$ (e.g., months). For each calendar date $s \leq t$, let $X_s^{(t)} \in \mathbb{R}^N$ denote the N-dimensional predictor vector dated s as recorded in vintage t.

Real-time vintage data. At forecast origin t, the forecaster observes the full vintage dataset

$$D_t \;:=\; \{X_s^{(t)} : s \leq t\}, \tag{1}$$

and the associated information set

$$\mathcal{I}_t^{RT} \;:=\; \sigma\Big(D_t,\ \{y_{j,s}\}_{s\leq t,\ j\leq J}\Big). \tag{2}$$

Final revised data. Let $X_s^{(\infty)}$ denote the fully revised (final) value for date s. The final-data environment uses the history

$$X_{\leq t}^{(\infty)} \;:=\; \{X_s^{(\infty)} : s \leq t\}, \tag{3}$$

with information set

$$\mathcal{I}_t^{F} \;:=\; \sigma\Big(X_{\leq t}^{(\infty)},\ \{y_{j,s}\}_{s\leq t,\ j\leq J}\Big). \tag{4}$$

Throughout, we treat the target series as final (i.e., $y_{j,t} \equiv y_{j,t}^{(\infty)}$) and allow revisions only through the predictor panel $\{X_s^{(t)}\}$ to isolate the mechanism of interest.[1] We use the environment index $e \in \{RT, F\}$ when a statement applies to both information sets.

2.2 Final-data factor structure and the revision process

Final-data factor model. The final revised predictor panel follows an approximate factor model:

$$X_t^{(\infty)} \;=\; \Lambda f_t + e_t, \qquad f_t \in \mathbb{R}^r,\ \Lambda \in \mathbb{R}^{N\times r},\ e_t \in \mathbb{R}^N, \tag{5}$$

[1] Allowing revisions in $y_{j,t}$ is straightforward and would introduce additional revision-induced terms later without altering the main logic.

where $r \ll N$ and e_t is idiosyncratic. For scale normalization, we adopt the standard convention

$$\frac{1}{N}\Lambda'\Lambda \;\to\; I_r \quad \text{as } N \to \infty, \tag{6}$$

and maintain the usual weak cross-sectional dependence conditions on e_t when needed.

Revision process. Real-time vintages differ from final data due to revisions:

$$X_s^{(t)} \;=\; X_s^{(\infty)} + R_{s|t}, \qquad s \le t, \tag{7}$$

where $R_{s|t} \in \mathbb{R}^N$ is the revision error for date s as of vintage t. We impose a minimal long-run consistency condition ensuring that final data are the long-run limit of the vintage process:

$$\lim_{m\to\infty} R_{s|s+m} \;=\; 0 \qquad \text{for each fixed } s. \tag{8}$$

Combining (5)–(7), the vintage panel admits the representation

$$X_s^{(t)} \;=\; \Lambda f_s + e_s + R_{s|t}, \qquad s \le t. \tag{9}$$

We interpret $R_{s|t}$ as a measurement disturbance induced by the revision process; no independence assumptions are imposed at this stage.

2.3 Targets

Let $\{y_{j,t}\}_{j=1}^J$ denote the target series used to define and aggregate uncertainty. Each target loads on the same latent factor:

$$y_{j,t} \;=\; a_j' f_t + \varepsilon_{j,t}, \qquad a_j \in \mathbb{R}^r, \quad j = 1, \ldots, J, \tag{10}$$

where $\varepsilon_{j,t}$ is series-specific. The common component $a_j' f_t$ is the channel through which aggregate shocks propagate into the targets.

2.4 Factor dynamics and multi-step propagation

Factor law of motion. The latent factor follows a stable VAR(q):

$$f_t \;=\; \Phi_1 f_{t-1} + \cdots + \Phi_q f_{t-q} + u_t, \qquad \mathbb{E}(u_t \mid \mathcal{G}_{t-1}) = 0, \qquad \mathrm{Var}(u_t \mid \mathcal{G}_{t-1}) = \Sigma_{u,t}, \tag{11}$$

where $\mathcal{G}_{t-1} := \sigma(g_{t-1}, g_{t-2}, \ldots)$ denotes the latent filtration generated by the factor state history, and $\Sigma_{u,t}$ is allowed to vary over time. Time variation in $\Sigma_{u,t}$ is the primitive source of *factor-innovation uncertainty* targeted by the factor-based index introduced in Section 3.

Companion form and propagation operator. Define the stacked state

$$g_t \ := \ (f_t', f_{t-1}', \ldots, f_{t-q+1}')' \in \mathbb{R}^{rq}, \tag{12}$$

and write the companion form

$$g_t \ = \ A g_{t-1} + \tilde{u}_t, \qquad \tilde{u}_t \ := \ (u_t', 0', \ldots, 0')'. \tag{13}$$

Let $\Pi := (I_r, 0, \ldots, 0)$ so that $f_t = \Pi g_t$. For any horizon $h \geq 1$, define the propagation operator

$$\Psi_h \ := \ \Pi A^h, \qquad \text{so that} \qquad \mathbb{E}(f_{t+h} \mid \mathcal{G}_t) \ = \ \Psi_h g_t. \tag{14}$$

This operator governs both multi-step forecasts and the propagation of time-t proxy errors into future forecast errors.

2.5 Factor proxies and proxy errors in real time

Factor extraction as a mapping. At time t, the econometrician extracts factor proxies from the available panel history. We represent the extraction procedure as a mapping $\mathcal{E}(\cdot)$. The leading empirical case is principal components applied to the relevant data matrix; the abstract notation is used to keep the theory agnostic to the particular extraction method.

Real-time factor proxies with vintage indexing. Applying $\mathcal{E}(\cdot)$ to the vintage dataset D_t produces a sequence of factor proxies for dates $s \leq t$. We write this explicitly as

$$\{\hat{f}^{RT}_{s|t}\}_{s \leq t} \ := \ \mathcal{E}(D_t), \tag{15}$$

where $\hat{f}^{RT}_{s|t}$ is the proxy for the factor dated s constructed using vintage-t data. The stacked proxy state at forecast origin t is then

$$\hat{g}^{RT}_t \ := \ (\hat{f}^{RT\prime}_{t|t}, \hat{f}^{RT\prime}_{t-1|t}, \ldots, \hat{f}^{RT\prime}_{t-q+1|t})' \in \mathbb{R}^{rq}. \tag{16}$$

Final-data factor proxies. Analogously, applying $\mathcal{E}(\cdot)$ to the fully revised history $X^{(\infty)}_{\leq t}$ yields

$$\{\hat{f}^{F}_{s|t}\}_{s \leq t} \ := \ \mathcal{E}(X^{(\infty)}_{\leq t}), \tag{17}$$

and we define

$$\hat{g}^{F}_t \ := \ (\hat{f}^{F\prime}_{t|t}, \hat{f}^{F\prime}_{t-1|t}, \ldots, \hat{f}^{F\prime}_{t-q+1|t})' \in \mathbb{R}^{rq}. \tag{18}$$

Proxy errors and conditional covariance. Define the environment-specific proxy error

$$\nu^e_t \ := \ \hat{g}^e_t - g_t, \qquad e \in \{RT, F\}, \tag{19}$$

and its conditional covariance under the corresponding information set

$$\Sigma^e_{\nu,t} \ := \ \mathrm{Var}(\nu^e_t \mid \mathcal{I}^e_t), \qquad e \in \{RT, F\}. \tag{20}$$

The key real-time friction is that ν^{RT}_t inherits revision-driven measurement noise from (9), whereas ν^F_t is the factor-extraction error that would arise under fully revised data.

2.6 Discussion: revisions → proxy errors → forecast errors

Equation (9) shows that the real-time predictor panel differs from its final counterpart by the revision disturbance $R_{s|t}$. Even when the same extraction procedure $\mathcal{E}(\cdot)$ is applied, this implies that the real-time proxy state $\hat{g}_t^{RT}$ generally differs from $\hat{g}_t^F$ and from the latent state g_t. The resulting proxy error ν_t^{RT} in (19) therefore combines conventional extraction noise and revision-induced measurement noise. Section 3 shows that ν_t^{RT} enters multi-step forecast errors through the propagation operator Ψ_h in (14), generating a revision-driven component in real-time forecast-error uncertainty measures.

3 Uncertainty Indices in Real Time: Revision Wedges and a Revision-Robust Factor-Based Index

This section studies how data revisions affect forecast-error-based uncertainty measures when the forecaster operates in real time using vintage data. Building on the real-time factor DGP in Section 2, we define y-based uncertainty indices under each environment $e \in \{RT, F\}$ with information sets $\mathcal{I}_t^{RT}$ and $\mathcal{I}_t^F$, derive an additive decomposition of the RT–F gap that isolates a nonnegative *revision wedge* attributable to revision-induced noise in factor proxies, and establish conditions under which this wedge attenuates with a large predictor cross-section N. Since N is finite in practice, we then introduce a *revision-robust factor-based index* that targets factor-innovation uncertainty directly and show that its RT–F gap does not load on revision-induced proxy noise. Finally, we provide diagnostic accounting ratios (component shares) and relate the proposed real-time factor-based index to the conventional final-data y-based benchmark, with a brief bridge to feasible plug-in estimators implemented in Section 4.

3.1 Definitions and preliminary decomposition

Fix a horizon $h \geq 1$ and an environment $e \in \{\mathrm{RT}, \mathrm{F}\}$. Recall from Section 2 that $f_t \in \mathbb{R}^r$ follows a stable VAR(q) with companion state $g_t \in \mathbb{R}^{rq}$, and that the multi-step propagation operator is $\Psi_h := \Pi A^h$ so that $\mathbb{E}(f_{t+h} \mid \mathcal{G}_t) = \Psi_h g_t$, where $\mathcal{G}_t := \sigma(g_t, g_{t-1}, \ldots)$.

Implemented factor-based forecasts. For each target series $j = 1, \ldots, J$, the structural relation is $y_{j,t} = a_j' f_t + \varepsilon_{j,t}$ (Section 2.3). In environment e, the econometrician forms a proxy state $\hat{g}_t^e$ and proxy error $\nu_t^e := \hat{g}_t^e - g_t$ (Section 2.5), and implements the h-step-ahead forecast

$$\hat{y}^e_{j,t+h|t} \ := \ a_j' \, \Psi_h \, \hat{g}_t^e, \qquad e \in \{\mathrm{RT}, \mathrm{F}\}. \tag{21}$$

The associated h-step forecast error is

$$e^e_{j,t}(h) \ := \ y_{j,t+h} - \hat{y}^e_{j,t+h|t}. \tag{22}$$

y-based uncertainty indices. Define the series-level conditional forecast-error uncertainty by

$$U_{j,t}^{y,e}(h)^2 \; := \; \mathrm{Var}\big(e_{j,t}^e(h) \mid \mathcal{I}_t^e\big), \qquad e \in \{\mathrm{RT}, \mathrm{F}\}, \tag{23}$$

and the macro aggregate (JLN-style) index as the cross-sectional average

$$U_t^{y,e}(h)^2 \; := \; \frac{1}{J} \sum_{j=1}^{J} U_{j,t}^{y,e}(h)^2, \qquad e \in \{\mathrm{RT}, \mathrm{F}\}. \tag{24}$$

Averaging matrix. Let

$$\bar{A} \; := \; \frac{1}{J} \sum_{j=1}^{J} a_j a_j' \; \in \; \mathbb{R}^{r \times r}, \tag{25}$$

which aggregates the quadratic exposure of targets to the common factor.

Factor forecast-error variance. Define the latent h-step factor forecast error and its conditional variance as

$$u_{t,h} \; := \; f_{t+h} - \mathbb{E}(f_{t+h} \mid \mathcal{G}_t) \; = \; f_{t+h} - \Psi_h g_t, \qquad \Sigma_{u,t}^{(h)} \; := \; \mathrm{Var}(u_{t,h} \mid \mathcal{G}_t). \tag{26}$$

Lemma 1 (Forecast-error representation)**.** *For each j and each environment $e \in \{\mathrm{RT}, \mathrm{F}\}$,*

$$e_{j,t}^e(h) \; = \; a_j' u_{t,h} \; + \; \varepsilon_{j,t+h} \; - \; a_j' \Psi_h \nu_t^e. \tag{27}$$

Assumption 1 (Conditional orthogonality as moment restrictions)**.** *Fix $h \geq 1$ and $e \in \{\mathrm{RT}, \mathrm{F}\}$. For each j, the components in* (27) *satisfy the conditional moment restrictions*

$$\mathbb{E}\big[u_{t,h}\, \varepsilon_{j,t+h} \mid \mathcal{I}_t^e\big] = 0, \tag{28}$$

$$\mathbb{E}\big[u_{t,h}\, (\nu_t^e)' \mid \mathcal{I}_t^e\big] = 0, \tag{29}$$

$$\mathbb{E}\big[\varepsilon_{j,t+h}\, (\nu_t^e)' \mid \mathcal{I}_t^e\big] = 0. \tag{30}$$

Remark 1 (Moment-restriction interpretation)**.** *Assumption 1 imposes conditional orthogonality in the form of the moment restrictions* (28)–(30)*. Economically, $u_{t,h}$ represents the common factor innovation over horizon h and is the primitive source of* factor-innovation uncertainty*; $\varepsilon_{j,t+h}$ captures series-specific news; and ν_t^e summarizes the factor-extraction error induced by using the information set $\mathcal{I}_t^e$ (including, in real time, vintage-data revisions). The restrictions state that, conditional on $\mathcal{I}_t^e$, (i) common factor innovations are orthogonal to series-specific news, (ii) common factor innovations are orthogonal to factor-extraction error, and (iii) series-specific news is orthogonal to factor-extraction error. Technically, these moment restrictions eliminate cross terms in the conditional second-moment expansion of the forecast error implied by* (27)*, so that the resulting uncertainty index admits a transparent additive decomposition.*

Lemma 2 (Additive decomposition of y-based uncertainty)**.** *Under Assumption 1, for each j and each environment $e \in \{\mathrm{RT}, \mathrm{F}\}$,*

$$U^{y,e}_{j,t}(h)^2 \;=\; \mathbb{E}\big[a_j'\Sigma^{(h)}_{u,t}a_j \mid \mathcal{I}^e_t\big] \;+\; \mathrm{Var}(\varepsilon_{j,t+h} \mid \mathcal{I}^e_t) \;+\; a_j'\Psi_h\,\Sigma^e_{\nu,t}\,\Psi_h' a_j, \tag{31}$$

where $\Sigma^e_{\nu,t} := \mathrm{Var}(\nu^e_t \mid \mathcal{I}^e_t)$ is defined in (20)*. Averaging over j yields the macro decomposition*

$$U^{y,e}_t(h)^2 \;=\; \mathbb{E}\Big[\mathrm{tr}\big(\bar{A}\,\Sigma^{(h)}_{u,t}\big)\,\Big|\,\mathcal{I}^e_t\Big] \;+\; \bar{\sigma}^{2,e}_{\varepsilon,t}(h) \;+\; \mathrm{tr}\big(\bar{A}\,\Psi_h\,\Sigma^e_{\nu,t}\,\Psi_h'\big), \tag{32}$$

where

$$\bar{\sigma}^{2,e}_{\varepsilon,t}(h) \;:=\; \frac{1}{J}\sum_{j=1}^{J}\mathrm{Var}(\varepsilon_{j,t+h} \mid \mathcal{I}^e_t). \tag{33}$$

Discussion. Equation (32) decomposes measured y-based uncertainty into a factor-innovation component, an idiosyncratic component, and a factor-extraction component. The next subsection compares the real-time and final-data versions of (32) and isolates a revision wedge that arises when real-time factor proxies inherit revision-induced measurement noise.

3.2 The RT–F gap, the revision wedge, and the remainder

This subsection compares the real-time and final-data versions of the y-based macro uncertainty index and isolates a distinct component attributable to revision-induced noise in factor proxies.

RT–F gap under the preliminary decomposition. By Lemma 2, for each $e \in \{\mathrm{RT}, \mathrm{F}\}$,

$$U^{y,e}_t(h)^2 \;=\; \mathbb{E}\Big[\mathrm{tr}\big(\bar{A}\,\Sigma^{(h)}_{u,t}\big)\,\Big|\,\mathcal{I}^e_t\Big] \;+\; \bar{\sigma}^{2,e}_{\varepsilon,t}(h) \;+\; \mathrm{tr}\big(\bar{A}\,\Psi_h\,\Sigma^e_{\nu,t}\,\Psi_h'\big),$$

where $\bar{\sigma}^{2,e}_{\varepsilon,t}(h)$ is defined in (33) and $\Sigma^e_{\nu,t} := \mathrm{Var}(\nu^e_t \mid \mathcal{I}^e_t)$ is defined in (20). Taking differences yields the identity

$$\begin{aligned} U^{y,\mathrm{RT}}_t(h)^2 - U^{y,\mathrm{F}}_t(h)^2 = {}& \underbrace{\mathbb{E}\Big[\mathrm{tr}\big(\bar{A}\,\Sigma^{(h)}_{u,t}\big)\,\Big|\,\mathcal{I}^{\mathrm{RT}}_t\Big] - \mathbb{E}\Big[\mathrm{tr}\big(\bar{A}\,\Sigma^{(h)}_{u,t}\big)\,\Big|\,\mathcal{I}^{\mathrm{F}}_t\Big]}_{\Delta CI_t(h)} \\ & + \underbrace{\bar{\sigma}^{2,\mathrm{RT}}_{\varepsilon,t}(h) - \bar{\sigma}^{2,\mathrm{F}}_{\varepsilon,t}(h)}_{\Delta\varepsilon_t(h)} + \mathrm{tr}\Big(\bar{A}\,\Psi_h\big(\Sigma^{\mathrm{RT}}_{\nu,t} - \Sigma^{\mathrm{F}}_{\nu,t}\big)\Psi_h'\Big). \end{aligned} \tag{34}$$

Assumption 2 (Factor-extraction error decomposition: revision vs. non-revision)**.** *For each t, in the real-time environment there exist random vectors $\nu^{\mathrm{RT}}_{id,t}$ and $\nu_{rev,t}$ such that the factor-extraction error admits the additive representation*

$$\nu^{\mathrm{RT}}_t \;=\; \nu^{\mathrm{RT}}_{id,t} \;+\; \nu_{rev,t}, \tag{35}$$

with the maintained conditional moment restriction

$$\mathbb{E}\big[\nu^{\mathrm{RT}}_{id,t}\,\nu'_{rev,t}\mid \mathcal{I}^{\mathrm{RT}}_t\big] \;=\; 0. \tag{36}$$

In the final-data environment, the factor-extraction error contains no revision-induced component, so $\nu^{\mathrm{F}}_t = \nu^{\mathrm{F}}_{id,t}$.

Lemma 3 (Covariance decomposition implied by Assumption 2)**.** *Under Assumption 2, the real-time factor-extraction-error covariance satisfies*

$$\Sigma^{\mathrm{RT}}_{\nu,t} \;=\; \Sigma^{\mathrm{RT}}_{\nu,id,t} \;+\; \Sigma_{\nu,rev,t}, \qquad \Sigma_{\nu,rev,t} \succeq 0, \tag{37}$$

where $\Sigma^{\mathrm{RT}}_{\nu,id,t} = \mathrm{Var}(\nu^{\mathrm{RT}}_{id,t}\mid \mathcal{I}^{\mathrm{RT}}_t)$ *and* $\Sigma_{\nu,rev,t} = \mathrm{Var}(\nu_{rev,t}\mid \mathcal{I}^{\mathrm{RT}}_t)$.

Remark 2 (Scope of Assumption 2)**.** *Assumption 2 is stated at the level of conditional moments. It does not impose a specific parametric model for data revisions nor a particular factor-extraction method. Instead, it requires only that, in real time, the factor-extraction error can be represented as the sum of a non-revision component and a revision-induced component that is conditionally orthogonal to the former. Lemma 3 then delivers the additive covariance decomposition* (37)*, which is the object used in our uncertainty-index accounting.*

Definition 1 (Revision wedge)**.** *For any horizon* $h \geq 1$*, define the revision wedge*

$$C^{rev}_t(h) \;:=\; \mathrm{tr}\big(\bar{A}\,\Psi_h\,\Sigma_{\nu,rev,t}\,\Psi'_h\big). \tag{38}$$

Proposition 1 (RT–F gap decomposition and the revision wedge)**.** *Fix any* $h \geq 1$*. Under Assumptions 1 and 2 (and hence Lemma 3), the real-time versus final-data gap admits the decomposition*

$$U^{y,\mathrm{RT}}_t(h)^2 - U^{y,\mathrm{F}}_t(h)^2 \;=\; C^{rev}_t(h) \;+\; R_t(h), \tag{39}$$

where the revision wedge *is*

$$C^{rev}_t(h) \;:=\; \mathrm{tr}\Big(\bar{A}\,\Psi_h\,\Sigma_{\nu,rev,t}\,\Psi'_h\Big) \;\geq\; 0, \tag{40}$$

with $\Sigma_{\nu,rev,t} = \mathrm{Var}(\nu_{rev,t}\mid \mathcal{I}^{\mathrm{RT}}_t)$ *as in Lemma 3. The remainder* $R_t(h)$ *collects all non-revision differences between the two environments:*

$$R_t(h) \;:=\; \Delta CI_t(h) \;+\; \Delta\varepsilon_t(h) \;+\; \Delta\nu_{id,t}(h), \tag{41}$$

where

$$\Delta\nu_{id,t}(h) \;:=\; \mathrm{tr}\Big(\bar{A}\,\Psi_h\big(\Sigma^{\mathrm{RT}}_{\nu,id,t} - \Sigma^{\mathrm{F}}_{\nu,id,t}\big)\Psi'_h\Big), \tag{42}$$

and $\Sigma^{\mathrm{F}}_{\nu,id,t} = \mathrm{Var}(\nu^{\mathrm{F}}_{id,t}\mid \mathcal{I}^{\mathrm{F}}_t)$ *(with* $\nu^{\mathrm{F}}_t = \nu^{\mathrm{F}}_{id,t}$ *by Assumption 2).*

Corollary 1 (Nonnegativity of the revision wedge)**.** *For any* $h \geq 1$, $C^{rev}_t(h) \geq 0$.

Interpretation. Proposition 1 isolates a distinct and nonnegative component, $C_t^{rev}(h)$, through which revisions inflate real-time y-based uncertainty. The remainder $R_t(h)$ captures all other differences between the real-time and final-data environments: (i) differences in conditioning sets that affect the conditional mean of factor-innovation uncertainty ($\Delta CI_t(h)$), (ii) differences in idiosyncratic uncertainty ($\Delta\varepsilon_t(h)$), and (iii) differences in non-revision proxy-error components ($\Delta\nu_{id,t}(h)$).

The next subsection studies when the revision wedge $C_t^{rev}(h)$ is asymptotically negligible as the predictor cross-section N grows, and clarifies why finite-N settings motivate a revision-robust factor-based construction.

3.3 Large-N attenuation of the revision wedge

Proposition 1 shows that real-time y-based uncertainty can be inflated by the revision wedge $C_t^{rev}(h)$. This subsection clarifies when that wedge becomes asymptotically negligible as the cross-sectional dimension N of the predictor panel grows.

A large-panel diversification condition. The key requirement is that the revision-induced component of the factor-extraction error is cross-sectionally diversified, so that its contribution to the conditional second moment of the factor-extraction error shrinks as the panel size grows.

Assumption 3 (Large-N diversification of revision-induced factor-extraction error)**.** *Fix a horizon $h \geq 1$. Let $\nu_{rev,t}$ be the revision-induced component in Assumption 2, and define $\Sigma_{\nu,rev,t} := \mathrm{Var}(\nu_{rev,t} \mid \mathcal{I}_t^{\mathrm{RT}})$. Suppose:*

1. ***Bounded aggregation and propagation.*** *The matrices $\bar{A}$ and Ψ_h have bounded operator norms:*

$$\|\bar{A}\|_{\mathrm{op}} \leq C_A, \qquad \|\Psi_h\|_{\mathrm{op}} \leq C_\Psi(h), \tag{43}$$

   for constants $C_A < \infty$ and $C_\Psi(h) < \infty$ that do not depend on N.

2. ***Revision shocks are not pervasive.*** *There exists a (possibly high-dimensional) vector of revision shocks r_t with conditional covariance $\Sigma_{r,t} := \mathrm{Var}(r_t \mid \mathcal{I}_t^{\mathrm{RT}})$ satisfying*

$$\|\Sigma_{r,t}\|_{\mathrm{op}} \leq C_r, \tag{44}$$

   uniformly in t.

3. ***Revision-induced extraction error is an averaged mapping of revision shocks.*** *There exists an $r \times \dim(r_t)$ matrix $G_{t,N}$ (measurable with respect to $\mathcal{I}_t^{\mathrm{RT}}$) such that*

$$\nu_{rev,t} = G_{t,N}\, r_t, \tag{45}$$

   and the mapping is cross-sectionally dispersed in the sense that

$$\|G_{t,N}\|_F^2 = O_p(N^{-1}), \tag{46}$$

   uniformly in t.

Assumption 3(i) is mild in our setting: $\bar{A}$ is an average of rank-one matrices $a_j a_j'$, and Ψ_h is determined by the stable factor dynamics in Section 2.4. The substantive content is Assumption 3(ii)–(iii), which formalizes the idea that revision shocks affect the extracted factors through dispersed cross-sectional weights, yielding diversification as N grows.

Lemma 4 (Rate for the revision-induced factor-extraction-error covariance)**.** *Under Assumption 3,*

$$\mathrm{tr}\big(\Sigma_{\nu,rev,t}\big) = O_p(N^{-1}), \tag{47}$$

uniformly in t.

Proposition 2 (Large-N attenuation of the revision wedge)**.** *Under Assumption 3, for any fixed $h \geq 1$,*

$$C_t^{rev}(h) \;=\; \mathrm{tr}\big(\bar{A}\,\Psi_h\,\Sigma_{\nu,rev,t}\,\Psi_h'\big) \;=\; O_p(N^{-1}). \tag{48}$$

Proof. Because $\Sigma_{\nu,rev,t} \succeq 0$ and $\bar{A} \succeq 0$,

$$C_t^{rev}(h) = \mathrm{tr}\big(\bar{A}\,\Psi_h\,\Sigma_{\nu,rev,t}\,\Psi_h'\big) \leq \|\bar{A}\|_{\mathrm{op}}\,\mathrm{tr}\big(\Psi_h\,\Sigma_{\nu,rev,t}\,\Psi_h'\big).$$

Moreover,

$$\mathrm{tr}\big(\Psi_h\,\Sigma_{\nu,rev,t}\,\Psi_h'\big) \leq \|\Psi_h\|_{\mathrm{op}}^2\,\mathrm{tr}\big(\Sigma_{\nu,rev,t}\big),$$

since $\mathrm{tr}(BMB') \leq \|B\|_{\mathrm{op}}^2 \mathrm{tr}(M)$ for $M \succeq 0$. Combining these inequalities with (43) and Lemma 4 yields (48). □

Remark 3 (Horizon damping under stable factor dynamics)**.** *If the factor dynamics are stable so that $\|A^h\|_{\mathrm{op}} \leq C\rho^h$ for some $\rho \in (0,1)$, then $\|\Psi_h\|_{\mathrm{op}} = \|\Pi A^h\|_{\mathrm{op}} \leq \|\Pi\|_{\mathrm{op}}\,C\rho^h$, and the bound in the proof of Proposition 2 implies the refined rate*

$$C_t^{rev}(h) \;=\; O_p\big(\rho^{2h} N^{-1}\big).$$

Thus, revision-driven inflation can be more pronounced at short horizons and is dampened at longer horizons when factor dynamics are sufficiently stable.

Implication. Proposition 2 shows that revision-driven inflation in real-time y-based uncertainty can be asymptotically negligible when the revision-induced component of the factor-extraction error is sufficiently diversified across series, so that the associated revision wedge shrinks at rate N^{-1}. Importantly, this attenuation is not automatic. If data revisions generate a non-diversified (i.e., pervasive) component in the revision-induced factor-extraction error—so that $\mathrm{tr}(\Sigma_{\nu,rev,t})$ fails to vanish—the revision wedge need not be small and may remain empirically relevant even in large panels.

The next subsection therefore introduces a revision-robust factor-based uncertainty index that targets factor-innovation uncertainty directly and, by construction, does not load on the revision-induced factor-extraction-error component.

3.4 Finite N: the proposed real-time factor index and its link to the JLN benchmark

Proposition 2 shows that the revision wedge $C_t^{rev}(h)$ may attenuate under strong diversification as $N \to \infty$. In practice, N is finite and the empirical objective is to construct a *real-time* uncertainty index that targets the same primitive object that underlies the conventional *final-data* benchmark used in the literature. In our notation, the JLN benchmark corresponds to the final-data y-based macro index $U_t^{y,\mathrm{F}}(h)^2$. Our proposed index is the *real-time factor-based index* defined by

$$U_t^{f,\mathrm{RT}}(h)^2 \;:=\; \mathbb{E}\Big[\operatorname{tr}\!\big(\bar{A}\,\Sigma_{u,t}^{(h)}\big)\,\Big|\,\mathcal{I}_t^{\mathrm{RT}}\Big], \tag{49}$$

which targets factor-innovation uncertainty directly.

Definition 2 (Factor-innovation uncertainty and the proposed real-time index)**.** *Fix $h \geq 1$. Define the latent factor-innovation uncertainty target*

$$U_t^{f}(h)^2 \;:=\; \operatorname{tr}\!\big(\bar{A}\,\Sigma_{u,t}^{(h)}\big), \tag{50}$$

and its environment-specific versions by conditioning on $\mathcal{I}_t^e$:

$$U_t^{f,e}(h)^2 \;:=\; \mathbb{E}\Big[U_t^{f}(h)^2\,\Big|\,\mathcal{I}_t^e\Big] \;=\; \mathbb{E}\Big[\operatorname{tr}\!\big(\bar{A}\,\Sigma_{u,t}^{(h)}\big)\,\Big|\,\mathcal{I}_t^e\Big], \qquad e \in \{\mathrm{RT}, \mathrm{F}\}. \tag{51}$$

In particular, the proposed index in (49) *is $U_t^{f,\mathrm{RT}}(h)^2$.*

The final-data counterpart $U_t^{f,\mathrm{F}}(h)^2$ is introduced only as a conceptual comparator to formalize the role of information sets and to connect to the RT–F gap results in Sections 3.2–3.3.

Relation to the y-based benchmark. Lemma 2 implies that the y-based index bundles factor-innovation uncertainty with an idiosyncratic component and a factor-extraction-error component. Substituting Definition 2 into that decomposition yields, for each environment $e \in \{\mathrm{RT}, \mathrm{F}\}$ and any $h \geq 1$,

$$U_t^{y,e}(h)^2 \;=\; U_t^{f,e}(h)^2 \;+\; \bar{\sigma}_{\varepsilon,t}^{2,e}(h) \;+\; \operatorname{tr}\!\big(\bar{A}\,\Psi_h\,\Sigma_{\nu,t}^{e}\,\Psi_h'\big), \tag{52}$$

where $\bar{\sigma}_{\varepsilon,t}^{2,e}(h)$ is defined in (33) and $\Sigma_{\nu,t}^{e} := \mathrm{Var}(\nu_t^e \mid \mathcal{I}_t^e)$ is defined in (20).

Mechanism and advantage. Equation (52) makes the mechanism transparent. Data revisions affect measured real-time y-based uncertainty only through the factor-extraction-error component $\operatorname{tr}\!\big(\bar{A}\,\Psi_h\,\Sigma_{\nu,t}^{\mathrm{RT}}\,\Psi_h'\big)$, since this is the only term that loads on the real-time factor-extraction error ν_t^{RT}. By contrast, the proposed index $U_t^{f,\mathrm{RT}}(h)^2$ is defined directly on the latent factor forecast-error variance $\Sigma_{u,t}^{(h)}$ and therefore targets factor-innovation uncertainty without loading on the factor-extraction-error covariance.

Proposition 3 (Benchmark gap: $U^{f,\mathrm{RT}}$ versus the JLN final-data benchmark $U^{y,\mathrm{F}}$)**.** *For any $h \geq 1$,*

$$U_t^{f,\mathrm{RT}}(h)^2 - U_t^{y,\mathrm{F}}(h)^2 \;=\; \Delta CI_t(h) \;-\; \bar{\sigma}^{2,\mathrm{F}}_{\varepsilon,t}(h) \;-\; \mathrm{tr}\big(\bar{A}\,\Psi_h\,\Sigma^{\mathrm{F}}_{\nu,t}\,\Psi_h'\big), \tag{53}$$

where $\Delta CI_t(h)$ is defined in (34)*. In particular,* (53) *contains no revision-induced factor-extraction-error component and therefore involves no revision wedge.*

Interpretation. Proposition 3 provides a transparent decomposition linking the proposed real-time factor-based index $U_t^{f,\mathrm{RT}}(h)^2$ to the JLN-style final-data benchmark $U_t^{y,\mathrm{F}}(h)^2$. Relative to $U_t^{f,\mathrm{RT}}(h)^2$, the benchmark differs only through (i) the information-set difference in factor-innovation uncertainty, $\Delta CI_t(h)$, and (ii) the non-common components embedded in the benchmark, namely idiosyncratic uncertainty $\bar{\sigma}^{2,\mathrm{F}}_{\varepsilon,t}(h)$ and the final-data factor-extraction-error component $\mathrm{tr}(\bar{A}\,\Psi_h\,\Sigma^{\mathrm{F}}_{\nu,t}\,\Psi_h')$. Importantly, no revision wedge appears in (53).

Corollary 2 (Revision-driven contamination: y-based versus factor-based RT–F gaps)**.** *For any $h \geq 1$,*

$$U_t^{f,\mathrm{RT}}(h)^2 - U_t^{f,\mathrm{F}}(h)^2 \;=\; \Delta CI_t(h), \tag{54}$$

so the RT–F gap of the proposed index contains no revision-driven component. Moreover, under Assumption 2 (and hence Lemma 3),

$$\Big(U_t^{y,\mathrm{RT}}(h)^2 - U_t^{y,\mathrm{F}}(h)^2\Big) - \Big(U_t^{f,\mathrm{RT}}(h)^2 - U_t^{f,\mathrm{F}}(h)^2\Big) \;=\; C_t^{rev}(h) + \Delta\varepsilon_t(h) + \Delta\nu_{id,t}(h), \tag{55}$$

where $C_t^{rev}(h)$ is the revision wedge in (40)*, and $\Delta\varepsilon_t(h)$ and $\Delta\nu_{id,t}(h)$ are defined in* (34) *and* (42)*.*

Implication for practice. Corollary 2 formalizes the main advantage of the proposed index. Unlike the real-time y-based index, whose RT–F gap loads on revision-induced factor-extraction error through the revision wedge $C_t^{rev}(h)$, the proposed index $U_t^{f,\mathrm{RT}}(h)^2$ targets factor-innovation uncertainty in real time without loading on the revision-induced factor-extraction-error component. Together with Proposition 3, this shows that $U_t^{f,\mathrm{RT}}(h)^2$ serves as a real-time proxy for the factor-innovation component underlying the JLN final-data benchmark $U_t^{y,\mathrm{F}}(h)^2$, while avoiding revision-driven inflation.

The next subsection introduces component-share accounting ratios, relates the proposed real-time factor-based index to the benchmark in empirical reporting, and provides a brief bridge to feasible plug-in estimators implemented in Section 4.

3.5 Empirical counterparts

The theoretical results in Section 3 deliver accounting relationships that (i) link the JLN-style final-data benchmark $U_t^{y,\mathrm{F}}(h)^2$ to the proposed real-time factor-based index $U_t^{f,\mathrm{RT}}(h)^2$ (Proposition 3), and (ii) characterize how revision-induced factor-extraction error enters real-time y-based uncertainty but is absent from the factor-based construction (Proposition 1 and Corollary 2). This subsection maps these relationships into a set of reporting quantities that are simple to compute and directly informative about the paper's main empirical claims.

Benchmark recovery: does $U^{f,\mathrm{RT}}$ track $U^{y,\mathrm{F}}$? Our benchmark follows Jurado, Ludvigson, and Ng (2015), which is implemented using fully revised (final) data. In our notation, this benchmark is $U_t^{y,\mathrm{F}}(h)^2$. Our proposed object is the real-time factor-based index $U_t^{f,\mathrm{RT}}(h)^2$ (Definition 2). For an initial visual comparison, we report standardized versions of the three time series

$$\left\{ U_t^{y,\mathrm{F}}(h)^2,\ U_t^{y,\mathrm{RT}}(h)^2,\ U_t^{f,\mathrm{RT}}(h)^2 \right\}, \tag{56}$$

where $U_t^{y,\mathrm{RT}}(h)^2$ serves as the natural naive real-time counterpart.[2]

To quantify the remaining discrepancy between the proposed real-time index and the JLN benchmark, we define the *benchmark gap*

$$B_t(h) \ := \ U_t^{f,\mathrm{RT}}(h)^2 - U_t^{y,\mathrm{F}}(h)^2. \tag{57}$$

Proposition 3 provides an exact decomposition of $B_t(h)$ and, in particular, implies that $B_t(h)$ contains no revision-wedge term. Empirically, we summarize benchmark recovery using: (i) comovement measures such as $\left(U_t^{f,\mathrm{RT}}(h)^2, U_t^{y,\mathrm{F}}(h)^2\right)$ and corresponding R^2 from a linear projection; (ii) dispersion measures for $\{B_t(h)\}$ (e.g., its mean, standard deviation, and RMSE); and (iii) calibration from the projection

$$U_t^{y,\mathrm{F}}(h)^2 = \alpha_h + \beta_h\, U_t^{f,\mathrm{RT}}(h)^2 + \eta_{t,h},$$

with the same projection using $U_t^{y,\mathrm{RT}}(h)^2$ in place of $U_t^{f,\mathrm{RT}}(h)^2$ as a comparison. Together, these summaries assess whether the proposed factor-based index provides a closer real-time proxy for the final-data benchmark than the naive real-time y-based implementation.

Revision-related component: isolating the non-common real-time wedge. A central distinction in the theory is between (i) information-set differences that affect conditional factor-innovation uncertainty and (ii) revision-induced factor-extraction error that enters implemented real-time y-based forecast-error volatility. Corollary 2 implies that the RT–F gap of the factor-based index equals the factor-innovation conditioning term $\Delta CI_t(h)$, whereas the RT–F gap of the y-based index additionally loads on non-common components, including the revision wedge. This motivates the following difference-in-gaps statistic:

$$G_t(h) \ := \ \Big(U_t^{y,\mathrm{RT}}(h)^2 - U_t^{y,\mathrm{F}}(h)^2\Big) - \Big(U_t^{f,\mathrm{RT}}(h)^2 - U_t^{f,\mathrm{F}}(h)^2\Big). \tag{58}$$

[2]Throughout, standardization refers to an affine rescaling applied separately to each series (e.g., demeaning and dividing by its sample standard deviation) to facilitate comparison of comovement.

By Corollary 2, $G_t(h)$ removes the factor-innovation conditioning difference and isolates the non-common real-time component:

$$G_t(h) \;=\; C_t^{rev}(h) \;+\; \Delta\varepsilon_t(h) \;+\; \Delta\nu_{id,t}(h), \tag{59}$$

where $C_t^{rev}(h) \geq 0$ is the revision wedge. We report $\{G_t(h)\}$ as a time-series summary of the magnitude and timing of non-common real-time distortions in the y-based index. For a scale-free measure, we report the ratio

$$\mathrm{InflShare}_t(h) \;:=\; \frac{G_t(h)}{U_t^{y,\mathrm{RT}}(h)^2}, \tag{60}$$

which measures the contribution of the non-common real-time component relative to the level of the real-time y-based index.[3]

Additional implications: horizon and large-N patterns. We also examine whether the empirical diagnostics exhibit the qualitative patterns implied by the model. First, under stable factor dynamics, revision-driven distortions propagated through Ψ_h are expected to be more pronounced at short horizons and dampened at longer horizons (Remark 3). Accordingly, we report how the dispersion of $G_t(h)$ varies with h (e.g., $h \in \{1, 3, 12\}$).

Second, Proposition 2 implies that the revision wedge can attenuate with the cross-sectional dimension under diversification. To assess this implication, we recompute the indices and $G_t(h)$ using nested predictor sets of increasing size N (or repeated subsampling within the available panel) and report how the variability of $G_t(h)$ evolves with N. These patterns are not formal tests; rather, they provide additional evidence on whether the attenuation mechanisms highlighted by the theory are quantitatively relevant in the data.

[3]Because $G_t(h)$ includes $\Delta\varepsilon_t(h)$ and $\Delta\nu_{id,t}(h)$, it need not be nonnegative at all dates. In empirical reporting we therefore summarize both the level and the dispersion of $\{G_t(h)\}$ (and, where useful, its positive part).