Nowcasting GDP using data revisions in an emerging economy*

Jason Cruz
jj.cruza@up.edu.pe

Diego Winkelried†
winkelried_dm@up.edu.pe

Javier Torres
j.torresgomez@up.edu.pe

School of Economics and Finance
Universidad del Pacífico (Lima, Perú)

This version: January 19, 2026

Abstract

This paper analyzes the revision process of monthly GDP growth in Peru using a newly constructed real-time dataset. A unified empirical framework is developed to test rationality by jointly examining serial and vintage correlations in revisions. The results reveal systematic and predictable revision patterns, primarily driven by benchmarking procedures. Motivated by these findings, a simple forecast-adjustment model is proposed that improves early assessments of economic activity by anticipating subsequent revisions. The evidence illustrates how revision dynamics in an emerging economy can be exploited to enhance nowcasting performance.

JEL Classification : C22, C82, C53, E01, O11, O54
Keywords : Data revisions, rationality tests, nowcasting, real-time data, emerging markets.

*This paper was presented at the 43rd Annual Conference of Economists of the Central Reserve Bank of Peru. We gratefully acknowledge the financial support of the Vice-Rectorate for Research at Universidad del Pacífico (8th Research Contest 2023), as well as funding from the Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) and the Programa Nacional de Investigación Científica y Estudios Avanzados (PROCIENCIA) under the call "E041-2025-04 Proyectos de Investigación en Ciencias Sociales". The authors alone are responsible for the views expressed and for any errors that may remain in this paper. This research uses publicly available, aggregate macroeconomic data and does not involve human subjects, experiments, or confidential information; therefore, no ethical approval was required. The authors declare no conflict of interest. The real-time GDP dataset constructed for this study, together with replication files, will be made available upon publication. All underlying data originate from publicly accessible statistical bulletins.

†Corresponding author. School of Economics and Finance, Universidad del Pacífico, Av. Salaverry 2020, Lima 15072, Peru. Phone: (+511) 219 0100 extension 2713.

1 Introduction

Macroeconomic data revisions are a routine feature of statistical practice. Aggregates such as GDP growth are often released with delay, under limited information, and subsequently updated as more comprehensive data become available (Lee, 2018). These revisions, while necessary, can have substantial implications. Academic research based on preliminary figures may yield different conclusions than research based on final values (Howrey, 1978; Stark and Croushore, 2002). From a policy standpoint, revisions affect how agents and institutions interpret the state of the economy and respond to it. They also shape the accuracy of nowcasting, where early estimates are used to gauge current conditions. Consequently, a rich literature has developed to study the properties of data revisions, assess their informational content, and explore whether they can be used to improve real-time inference (Croushore, 2011).

Much of this literature has centered on the experience of industrialized countries, for which real-time datasets (RTDs) are readily available. In contrast, studies on emerging economies are scarce, often due to the lack of systematic vintage repositories (cf. Yadav, 2021). Many statistical agencies in these countries do not retain historical versions of their estimates, replacing preliminary values with revised ones without archiving prior releases. As a result, researchers interested in these economies must first build their own RTDs, a task made difficult by the reliance on hard-copy bulletins, scanned reports, or other formats not suited for automated processing (Kishor, 2011; Ramos et al., 2004). Advances in optical character recognition and related technologies have recently made this task more feasible. In this paper, we follow such an approach to construct a RTD for monthly GDP growth in Peru, aiming to fill a gap in the literature and contribute a usable framework for similar efforts in other emerging markets.

Once a RTD is available, the primary challenge is determining whether preliminary releases can be improved in real time. To this end, a key intermediate step is to assess whether revisions are rational. By rationality, we mean that each release should make efficient use of the information available at the time of publication, such that subsequent updates cannot be systematically predicted (Mankiw et al., 1984; Nordhaus, 1987). To assess this, we adopt a widely used regression-based framework for testing rationality (Clements, 1997; Ducoudré et al., 2020), and propose a unified omnibus regression that nests several well-known tests as special cases. This unified approach not only allows us to determine whether a particular release fails rationality but also helps identify the source of the failure: bias, serial correlation across time, or cross-release dependence within a given month. Moreover, our framework extends naturally to study institutional mechanisms, such as benchmarking practices, which attempt to reconcile high-frequency series (e.g., monthly production) with more reliable lower-frequency aggregates (e.g., quarterly or annual totals). These procedures are known to generate serial and cross-release dependence in revisions (Durbin and Quenneville, 1997; Lee, 2018), which can be captured within our specification.

A rejection of rationality implies that available information at the time of the release was not fully exploited, opening the door for real-time improvements. This idea is foundational in the growing nowcasting literature (Bańbura et al., 2013). Our paper contributes to this agenda by placing forecasting considerations at the center of the analysis, proposing a simple method to adjust early releases based on the same regression framework used for the rationality tests. The resulting nowcasting equation leads to an analytical correction that brings preliminary values closer to the final figure. In contrast to more complex filtering approaches based on state-space models (e.g., Jacobs and van Norden, 2011; Aruoba et al., 2016), the adjustment is transparent, easily reproducible, and uses only information from past releases and revisions, making it suitable for practical forecasting applications.

The remainder of the paper is structured as follows. Section 2 reviews the literature on data revisions, with emphasis on rationality, benchmarking, and nowcasting. Section 3 presents the methodological framework. Section 4 describes the data collection process and the newly assembled dataset on monthly GDP growth in Peru, discusses the results of the rationality tests, and applies the proposed adjustment method. We find compelling evidence of irrationality in early releases, which tend to understate the final value and display both serial and vintage correlation. These patterns are consistent with benchmarking procedures that smooth monthly updates around quarterly reference points. Moreover, the nowcasting method yields statistically and economically significant improvements in prediction accuracy, highlighting the forecasting value of exploiting revision dynamics. Section 5 concludes. An online supplement provides additional details on data construction and robustness checks.

2 Literature Review

Data revisions are an inevitable feature of the production of official economic statistics. Whenever a statistical agency publishes an aggregate and faces even minor difficulties in collecting or processing the available information, preliminary figures may later be revised (Lee, 2018). Early contributions document that such revisions arise both from practical limitations in data collection and from institutional decisions about which inputs to incorporate at each stage (Croushore, 2006). Thus, revisions should not be interpreted solely as errors, but as a natural mechanism through which agencies adjust their estimates in light of new information (Croushore, 2011). Studies such as Stark (2011), Faust et al. (2005), and Kapetanios and Yates (2010) suggest that revisions may even be optimal under certain informational efficiency assumptions.

Revisions are not innocuous and can have important academic and policy implications. The early study of Howrey (1978) shows that econometric models are often estimated with preliminary data, which differ systematically from final revised values. Ignoring this difference biases parameter estimates and degrades forecast performance. More recently, Stark and Croushore (2002) demonstrated that prediction errors appear deceptively smaller when using revised data. This discussion highlights the importance of data vintages for proper evaluation. A key contribution in this regard is Croushore and Stark (2001, 2003), who introduced a comprehensive real-time data set for the United States, which has since become a cornerstone of the literature. Similar initiatives were developed for other advanced economies, such as Egginton et al. (2002) and Garratt and Vahey (2006) for the United Kingdom, Bernhardsen et al. (2005) for Norway, Bernhard (2016) for Switzerland, Fernandez et al. (2011) for the OECD, and Giannone et al. (2012) for the Euro area.[1]

At the policy level, Orphanides (2001) and Kugler et al. (2005) show that monetary policy assessments based on real-time data differ drastically from those using revised data, affecting the credibility of forecasts and policy. Aoki (2003) formalizes how monetary policy should respond to noisy indicators, highlighting that revisions are not only statistical but also a policy design challenge. Kishor (2011) provides a more recent account. Meanwhile, Amir-

[1] Most of this evidence comes from developed economies, with limited contributions for emerging economies. Sim et al. (2009) and Yadav (2021) are remarkable exceptions.

Ahmadi et al. (2017) provide empirical evidence that policymakers relying on real-time data underestimate the true effects of policy shocks, underscoring the risks of data revision. Similar concerns arise in fiscal policy: Cimadomo (2016) argues that the fiscal stance appears more counter-cyclical in real time, with political cycles and weak institutions influencing revisions. Castro et al. (2013) conclude that revisions have important implications for compliance with fiscal rules.

In a fully rational world, with both the statistical agency and economic agents using all available information efficiently, macroeconomic data revisions would be unpredictable and would only reflect the arrival of new information. This idea underlies the "news" hypothesis, seminally formulated by Mankiw et al. (1984) and Mankiw and Shapiro (1986), which holds that revisions lack systematic components and merely correct efficient forecast errors. This contrasts with the "noise" hypothesis, according to which preliminary figures contain measurement errors that revisions correct (Croushore, 2006, 2011). Early studies treat the news/noise distinction as the analytical benchmark, framing subsequent discussions of rationality and efficiency. For instance, Mork (1987) and DeJong (1987) analyze mixed models that combine measurement and forecast errors, providing the canonical benchmark for the idea of "well-behaved" revisions (unbiased, of low variance, and unpredictable) as a normative (though rarely observed) standard, while Croushore and Stark (2001) emphasize that revisions may embody either news or noise.

Subsequent research progressively qualifies and extends this benchmark, concluding that revisions in practice do not conform neatly to either the news or the noise hypothesis, nor to the notion of well-behaved revisions. For instance, Faust et al. (2005) and Aruoba (2008) argue that the news/noise hypotheses are not exhaustive, showing through predictability exercises that revisions can be biased, large, and partially anticipated. Hecq et al. (2019) reach a similar conclusion. Strohsal and Wolf (2020) provide further evidence of systematic patterns in revisions and demonstrate that filtering methods significantly enhance the accuracy of early estimates relative to the statistical agency's releases. Thus, as stated in Croushore (2006, 2011), empirical evidence points to a richer and more complex reality shaped by institutional practices, data collection procedures, and forecasting behavior, leading the field away from a rigid news/noise dichotomy and toward broader conceptual frameworks.

A central aspect of the literature is how to operationalize rationality empirically. The seminal work of Nordhaus (1987) proposes a notion of weak efficiency, according to which revisions should not display serial correlation if they only reflect new information. His finding of persistent positive correlations is attributed to gradual information incorporation (due to reputation or institutional factors), suggesting non-optimal behavior by the agency. In contrast, Clements (1997) find negative autocorrelation in revisions, interpreting them more as corrections of noise than as the incorporation of novel information. Bakhshi et al. (2005) also document similar autocorrelation within a different methodological framework.

A further explanation for correlation patterns in errors and revisions lies in the benchmarking procedures used by statistical agencies to ensure consistency between higher-frequency series and more reliable lower-frequency aggregates. Such procedures mechanically induce dependence between early estimates and later revisions, as noted by Lee (2018, ch. 6). The issue was formalized by Durbin and Quenneville (1997), who proposed state-space solutions to the benchmarking problem, and revisited by Weigand et al. (2018), who construct timely statistics from multiple data sources while addressing survey errors, breaks, differing sampling frequencies, and irregular observations. These contributions suggest that benchmarking may be a major source of apparent non-rationality, since reconciling surveys of different frequencies implies larger informational shocks in certain months, helping account for the dependence patterns observed in revisions.

The most common strategy to test for rationality is the regression framework of Mincer and Zarnowitz (1969), where prediction errors are regressed on variables that should, under rationality, be uncorrelated with them, with extensions in recent contributions such as Ducoudré et al. (2020). In the same spirit, rationality has often been assessed through autocorrelation tests of revisions or prediction errors. One strand emphasizes serial correlation across time for a given horizon, following the tradition of Mankiw et al. (1984), further developed by Mork (1987), Nordhaus (1987), and later by Aruoba (2008), Bernhard (2016), and Jacobs and van Norden (2016). A second strand emphasizes correlation across horizons for a given target event, which directly embeds the revision horizon into efficiency tests. This approach is less common but has gained traction in studies such as Clements (1997), Patton and Timmermann (2012), Clements (2022), and Strohsal and Wolf (2020), where revision properties are evaluated conditional on the horizon.

A third, less developed strand explicitly seeks to combine both dimensions, producing hybrid analyses. Examples include Hecq et al. (2019), who test for news and noise in multiple-revision processes, as well as empirical approaches that account simultaneously for sequential dynamics and horizon effects, such as in Garratt et al. (2008), Jacobs and van Norden (2011), and Aruoba et al. (2016). Nevertheless, this hybrid strand remains far less extensive than the other two, and systematic regression-based analyses of explicit time–horizon interactions are largely absent.

On the other hand, the recognition that revisions are not fully rational has led researchers to model their systematic patterns explicitly and to use them to adjust official preliminary estimates. The dominant approach is based on state-space modeling: early contributions include Conrad and Corrado (1979), Bordignon and Trivellato (1989) and Patterson (1995) while Jacobs and van Norden (2011), Cunningham et al. (2012), Aruoba et al. (2016), and Jacobs et al. (2022) provide more recent analyses. Complementary approaches include ARIMA modeling (Howrey, 1978; Guerrero, 1993), VAR forecasting (Kishor and Koenig, 2012), and error-correction models (Patterson, 2003; Hecq et al., 2019). See also Strohsal and Wolf (2020). Taken together, these approaches highlight that preliminary estimates are often suboptimal, which in turn motivates econometricians to improve predictions by exploiting real-time information. This perspective underlies the nowcasting literature, based on the use of high-frequency indicators to anticipate macroeconomic variables, notably GDP, before official release. See Bańbura et al. (2013) for a comprehensive review.

3 Methodology

This section presents the methodological framework used to evaluate the rationality of preliminary statistical releases. In particular, we test whether preliminary values are rational in the sense that they cannot be systematically improved using information available at the time of their publication. In addition to these tests, we also develop a nowcasting approach that is derived from the same framework, providing a novel and simple way to improve preliminary releases as forecasts of the final value.

The true value of the variable of interest is denoted by y_t. The revision process involves a sequence of preliminary predictions. Define y_t^h as the h-th prediction of y_t (i.e., the h-th release) for $h = 1, 2, \ldots, H$.[2] We assume that this process ends after H releases, so that the final release is, by construction, $y_t^H \equiv y_t$.

3.1 Revisions and prediction errors

We define the *revision* in the release h as:

$$r_t^h = y_t^h - y_t^{h-1}, \quad h = 2, 3, \ldots, H\,,$$

as the change between two consecutive predictions of the variable, that measures directly how new information is incorporated across successive releases. An related concept is the *prediction error* at release h:

$$e_t^h = y_t - y_t^h\,,$$

which is the difference between the true (final) value and the preliminary predictor y_t^h. The prediction errors are related to the revisions through:

$$r_t^h = e_t^{h-1} - e_t^h, \qquad e_t^h = \sum_{k=h+1}^{H} r_t^k, \quad h = 1, \ldots, H-1. \tag{1}$$

By construction, $e_t^H = 0$.

Our preferred measure to test for rationality is the revision, i.e., the incremental information that the statistical agency decides to add between two releases (cf. DeJong, 1987). In contrast, prediction errors are only observed *ex post*, once the final value is available, and they conflate all previous revisions into a single "distance from the truth" (cf. Bakhshi et al., 2005)." Rationality requires the flow of information – the revisions – to satisfy certain conditions (unbiasedness, unpredictability, etc.), whereas the conditions for prediction errors – a stock of future revisions – must be deduced from the properties of that flow. Thus, revisions are the natural object of study because they directly measure the information content of each release allowing a precise set of rationality tests (cf. Nordhaus, 1987).

However, we also present results on prediction errors because they remain popular in the literature (cf. Patton and Timmermann, 2012; Ducoudré et al., 2020), and because it is easier to apply the adjustments suggested to the rationality tests in order to achieve improved predictions through the analysis of prediction errors, as we discuss in Section 3.4. Prediction errors can be useful for evaluating the overall performance of releases, but not necessarily for diagnosing the revision mechanism itself.

3.2 Rationality based on revisions

A prediction is said to be *rational* if it uses all information available at the time. Formally, rationality requires that:

$$E(r_t^h \mid \mathcal{I}_t^h) = 0,$$

which means that the revision from release $h-1$ to release h cannot be systematically predicted from the information set at the time before the h-th release, $\mathcal{I}_t^h$.

[2] In our dataset, the first release of y_t is consistently published with a two-month delay (i.e, in month $t+2$). Subsequent releases appear monthly. Hence, predictions are indexed by h, the order of releases, rather than elapsed time since t $(t+h+1)$.

Since the information set $\mathcal{I}_t^h$ is unobservable, we test necessary conditions for rationality by checking whether revisions correlate with observable proxies $z_t^h \in \mathcal{I}_t^h$. In practice, this is implemented in a linear regression setting, where r_t^h is regressed on z_t^h, and possibly a constant. Rationality requires that the coefficients on these proxies be zero. Otherwise, if revisions are systematically related to z_t^h, we can conclude that original prediction was not fully informed, and thus not rational.

From now on, let ε_t denote a generic regression error. We use the parameters $(\alpha, \theta, \gamma, \rho)$ to represent the various conditions of rationality, with the important caveat that, despite the common notation, these parameters may not be comparable across equations. We nevertheless employ the same set of symbols in order to avoid cluttering the notation.

3.2.1 *Unbiasedness*

When $z_t^h = 1$, the regression is:

$$r_t^h = \alpha + \varepsilon_t \,. \tag{2}$$

The null hypothesis $H_0 : \alpha = 0$ states that revisions are unbiased, $\mathrm{E}(r_t^h) = 0$. Rejection of this null implies that revisions are systematically positive or negative, which indicates that preliminary predictions tend to underestimate or overestimate the final value.

3.2.2 *Time-series autocorrelation*

When $z_t^h = r_{t-1}^h$, the regression is:

$$r_t^h = \alpha + \rho\, r_{t-1}^h + \varepsilon_t \,. \tag{3}$$

The null hypothesis $H_0 : \rho = 0$ states that revisions are uncorrelated over time. If the data revision process is rational, each revision uses all available information and thus past revision patterns shouldn't help predict the current revision. Rejection means that revisions at a given release h are predictable from past revisions across time. Positive autocorrelation suggests gradual under-adjustment, while negative autocorrelation suggests overreaction followed by correction. It may indicate mechanical issues (like smoothing procedures), data delays, or institutional inertia in the updating process.

3.2.3 *Cross-release correlation*

When $z_t^h = r_t^{h-1}$, the regression is:

$$r_t^h = \alpha + \gamma\, r_t^{h-1} + \varepsilon_t \,. \tag{4}$$

The null hypothesis $H_0 : \gamma = 0$ requires that revisions be uncorrelated to the immediately preceding revision within the same target event. Rational predictions should use all available information as soon as it becomes available, which means that a large revision in one release should not be followed by another large revision or correction for further releases. Rejection implies that revisions follow a predictable pattern across releases, suggesting that new information is not fully incorporated in each release but instead spreads gradually across multiple updates.

3.2.4 Omnibus test

To capture both types of predictability jointly, we consider the augmented regression:

$$r_t^h = \alpha + \gamma\, r_t^{h-1} + \rho\, r_{t-1}^h + \varepsilon_t\,. \tag{5}$$

The null hypothesis $H_0 : \gamma = \rho = 0$ jointly assesses rationality across both dimensions. The term r_{t-1}^h tests correlation across time at a fixed release index (the t-dimension), while the term r_t^{h-1} tests correlation across releases within a fixed event (the h-dimension). Testing them jointly provides a compact assessment of whether revisions contain any systematic structure that contradicts rationality.

3.3 Rationality based on prediction errors

Since prediction errors accumulate "future" revisions, see equation (1), rationality also requires that $E(e_t^h \mid \mathcal{I}_t^h) = 0$, so that rational prediction errors are orthogonal to all information available at the time. As with revisions, we test necessary conditions by checking, also within a linear regression setting, if e_t^h correlates with observable proxies z_t^h.

3.3.1 Unbiasedness

When $z_t^h = 1$, the regression is:

$$e_t^h = \alpha + \varepsilon_t \tag{6}$$

The null hypothesis $H_0 : \alpha = 0$ asserts that preliminary releases are unbiased, $\mathrm{E}(e_t^h) = 0$. Rejection means that prediction errors are systematically positive or negative, contradicting rationality.

3.3.2 Mincer-Zarnowitz regression

Following the seminal work of Mincer and Zarnowitz (1969), consider a regression that relates the final realized value y_t to the prediction y_t^h: $y_t = \alpha + \beta y_t^h + \varepsilon_t$. Under $H_0 : \alpha = 0$ and $\beta = 1$, the h-th prediction is an unbiased (i.e., rational) predictor of y_t.

Subtracting y_t^h to both sides of the Mincer and Zarnowitz equation renders our testing equation with $z_t^h = y_t^h$:

$$e_t^h = \alpha + \theta\, y_t^h + \varepsilon_t\,. \tag{7}$$

The null hypotheses is rewritten as $H_0 : \alpha = \theta = 0$. Rejection indicates that prediction errors are correlated with the prediction itself, providing evidence against rationality.

3.3.3 cross-release correlation

One may also set $z_t^h = r_t^h$, the revision from release $h-1$ to h. This yields the regression:

$$e_t^h = \alpha + \gamma\, r_t^h + \varepsilon_t\,. \tag{8}$$

The null hypothesis $H_0 : \gamma = 0$ states that prediction errors do not correlate with revisions. The prediction error at release h reflects all future revisions still to come, so we inquire whether earlier revisions help explain the total amount of correction that remains to be made. If so, this suggests that the revision process is gradual or under-adjusting, as one revision is not enough to fully eliminate the prior error. Thus, (8) is closely related to (4) and so $H_0 : \gamma = 0$ can be interpreted as a cross-release correlation test.[3]

Interestingly, equation (8) can also be derived as an *encompassing test*, in the tradition of Hendry and Clements (2004) and Clements and Harvey (2010), to evaluate whether a prediction fully incorporates the information already contained in an earlier release. For two consecutive preliminary predictions y_t^{h-1} and y_t^h, define the bias-adjusted combined prediction as $y_t^c = -\alpha + \gamma y_t^{h-1} + (1-\gamma)y_t^h$. Subtracting the true value y_t from both sides gives the expression in terms of prediction errors $e_t^c = \alpha + \gamma e_t^{h-1} + (1-\gamma)e_t^h$. Rearranging yields:

$$e_t^h = \alpha + \gamma(e_t^{h-1} - e_t^h) + e_t^c \,,$$

which is a regression of e_t^h on the difference $e_t^{h-1} - e_t^h$, with e_t^c as the error term. Upon recalling that $r_t^h = e_t^{h-1} - e_t^h$, we get (8). This formulation allows us to unveil that under the null hypothesis $H_0 : \gamma = 0$, the prediction y_t^h "encompasses" y_t^{h-1}; i.e., the information contained in y_t^h but not in y_t^{h-1} (i.e., the revision $r_t^h = y_t^h - y_t^{h-1}$) does not provide any opportunity to reduce the size of the prediction error e_t^h. If such a null hypothesis holds, combining the two projections is unnecessary, given y_t^h.

3.3.4 Omnibus test

Finally, the above discussion highlights how encompassing tests can be embedded within the unified framework of our omnibus test, where cross-release and time-series dependencies are jointly assessed. The specification that nests all these uncorrelatedness conditions into a single regression is:

$$e_t^h = \alpha + \theta\, y_t^h + \gamma\, r_t^h + \rho\, r_{t-1}^h + \varepsilon_t \,. \tag{9}$$

The null hypothesis $H_0 : \alpha = \theta = \gamma = \rho = 0$ allows for a joint assessment of rationality. Rejection indicates that prediction errors display systematic structure, violating rationality.

3.4 Nowcasting

Consider equation (9). Even under rationality, the error term of this regression is likely to display serial correlation. For this reason, rationality should be tested using inference procedures that are robust to such correlation. For the purposes of prediction, however, it is more convenient to parameterize the autocorrelation and use the following variant:

$$e_t^h = \boldsymbol{\beta}' \boldsymbol{z}_t^h + \delta\, e_{t-1}^h + \varepsilon_t \,, \tag{10}$$

where $\boldsymbol{\beta} = (\alpha, \theta, \gamma, \rho)'$ and $\boldsymbol{z}_t^h = (1, y_t^h, r_t^h, r_{t-1}^h)'$. In practice, the autoregressive coefficient δ absorbs most of the error autocorrelation, so the residual ε_t in (10) can be expected to behave as white noise.

If rationality is rejected, then $\boldsymbol{\beta} \neq \boldsymbol{0}$. This means that the information available at the time of the releases is not fully exploited, implying that there is scope for improving preliminary releases as unbiased forecasts of the final value.

[3] If we begin with $e_t^h = \delta_0 + \delta_1 e_t^{h-1} + \varepsilon_t$ and use $e^{h-1} = r_t^h + e_t^h$, we obtain (8) for $\alpha = \delta_0/(1-\delta_1)$ and $\gamma = \delta_1/(1-\delta_1)$. Thus, $\gamma = 0$ if and only if $\delta_1 = 0$.

The objective is to nowcast $y_t = y_t^h + e_t^h$ using only the information available once the release y_t^h has been published. This information includes $\boldsymbol{z}_t^h$ and its history, **but not** e_{t-1}^h since the prediction error embeds "future" revisions. Thus, although equation (10) can be used to estimate $\boldsymbol{\beta}$ and δ historically, it cannot be applied directly for real-time forecasting.

In the spirit of Howrey (1978), Guerrero (1993) and Strohsal and Wolf (2020), a way to circumvent this limitation is to solve for the implicit expectation of (10), the "filtered" prediction. Given the structure of (10), this model-based expectation has a neat form that can be computed analytically.

Let $\mathcal{Z}_t^h = \{\boldsymbol{z}_{t-i}^h\}_{i=0}^{\infty} \subset \mathcal{I}_t^h$, and define $\hat{e}_t^h = \mathrm{E}(e_t^h \mid \mathcal{Z}_t^h)$. Note that $\mathrm{E}(\boldsymbol{z}_t^h \mid \mathcal{Z}_t^h) = \boldsymbol{z}_t^h$ and assume that $\mathrm{E}(\varepsilon_t \mid \mathcal{Z}_t^h) = 0$. Taking expectations conditional on $\mathcal{Z}_t^h$ in (10), we obtain:

$$\hat{e}_t^h - \delta\, \hat{e}_{t-1}^h = \boldsymbol{\beta}' \boldsymbol{z}_t^h, \quad \text{which renders} \quad \hat{e}_t^h = \boldsymbol{\beta}' \sum_{i=0}^{\infty} \delta^i \boldsymbol{z}_{t-i}^h\,. \tag{11}$$

The solution to this first-order difference equation is standard textbook material. If we denote by $\boldsymbol{Z}_t^h(\delta)$ the exponentially weighted moving sum of $\boldsymbol{z}_t^h$ and its history, it can be computed recursively as:

$$\boldsymbol{Z}_t^h(\delta) = \delta \boldsymbol{Z}_{t-1}^h(\delta) + \boldsymbol{z}_t^h\,,$$

which is readily available at the time of release h.

Thus, the nowcast of y_t associated with release h is given by the adjusted prediction:

$$\hat{y}_t = y_t^h + \hat{e}_t^h = y_t^h + \boldsymbol{\beta}' \boldsymbol{Z}_t^h(\delta)\,. \tag{12}$$

The nowcasting procedure described above uses information that could, in principle, have been incorporated into the releases themselves in order to improve their accuracy. As successive releases $\{y_t^h\}$ converge to the true value y_t, this procedure accelerates the sequence, effectively bringing early preliminary predictions closer to the final releases.

4 Results

4.1 Data collection

In Peru, the statistical agency – the National Institute of Statistics and Informatics (INEI) – produces a Monthly Index of National Production (INEI, 2016, 2021), whose publication is a highly anticipated event given its pivotal role in shaping economic decisions. This index is commonly referred to, even officially, as the "monthly GDP", and its monthly evolution is consistent with quarterly and annual GDP from the national accounts system. The first release, however, appears with a lag of around two months after the close of the reference month, as the INEI faces coverage limitations in processing information. Even once published, the data remain subject to a series of revisions. Although these revisions are formally mandated by law and arise for various reasons – such as delays in data collection, the complexity of processing sectoral sources, adjustments based on surveys, and methodological updates – the process has historically received little attention.

Chief Resolution No. 316-2003-INEI outlines the periodic revision process for sectoral economic sources. In practice, however, the revision process is largely undocumented, and agents are often unaware of it, leading preliminary figures to be treated as definitive. This lack of visibility is probably the main reason for the absence of a real-time dataset (RTD) in the country. The INEI does not maintain a repository of historical vintages, nor does it provide them through its website. The only consistent official source that allows revisions to be tracked over time is the Weekly Report (WR) of the Central Reserve Bank of Peru (BCRP), which republishes INEI figures and monthly updates in a coherent format that changes little over time.

To address these gaps and enhance transparency, this study constructs the RTD of monthly GDP year-on-year growth rates for Peru. The dataset was assembled from WR publications from 2000 to 2024 and is fully reproducible with automatic updating through our replication codes.[4]

The construction of the RTD followed a two-phase strategy, reflecting the availability and format of sources. Since 2013, the WR has been published online as digital (semi-structured) PDFs. For this period, we implemented automated extraction routines in Python, designed to handle heterogeneous formatting across documents. For the pre-2013 period, only printed bulletins were available. The process of extracting and processing the data presented significant challenges, particularly in relation to the quality of the source documents. The primary difficulty was the poor quality of the scanned documents, which varied in legibility. To convert these scans into usable data, we applied Optical Character Recognition (OCR) using the Tesseract OCR algorithm. While OCR facilitated the conversion of scanned documents into text, it was not always accurate and required significant manual intervention. Encoders were employed to correct errors introduced during the OCR process, ensuring the final dataset was accurate and reliable. The resulting RTD was developed under strict quality-control standards and designed for automatic updates. We document the data collection process in detail in an online supplement to the paper, and replication codes are also available.

To the best of our knowledge, there is very limited prior work constructing RTD for emerging economies. Some exceptions can be found in the literature on GDP revisions, where researchers built their own datasets, often by harvesting periodic publications not available in easily processed form – as our approach in this study. For example, Kishor (2011) compiled GDP vintages from the Reserve Bank of India's weekly statistical supplements, while Ramos et al. (2004) collected data from hard copies of press releases issued by the Brazilian statistical agency (IBGE). These cases underscore how scarce, yet valuable, RTD are in the context of emerging economies. We hope that our procedure and experience in assembling the Peruvian RTD can serve as a useful reference for future efforts.

4.2 Descriptive statistics

Figure 1 depicts the evolution of monthly GDP growth estimates across releases, from January 2001 to December 2024. The red line shows the first release of each monthly observation, while the blue lines trace the subsequent releases that incorporate progressively richer information sets. For clarity, only releases corresponding to even months are displayed. To ensure comparability, revisions during 2013 are excluded, as that year involved a change in the base year, which would otherwise confound regular revisions with methodological changes. Similarly, the period 2020–2021 is omitted due to the exceptional volatility generated by the introduction and subsequent lifting of COVID-19 lockdowns. The figure highlights that initial releases appear relatively volatile, reflecting the limited information available at the time of

[4] An online supplement provides detailed information on the data sources, collection procedures, processing steps, and terminology used to construct the RTD, focusing on the methodological workflow, quality-control standards, and replication resources.

publication, whereas later releases revise these values in light of more complete data. Although this smoother behavior of the final estimates is not easy to discern visually, it will be confirmed through formal statistical analysis in subsequent sections.

Figure 2 summarizes the distribution of revisions and prediction errors and across successive GDP releases. Prediction errors are wide and dispersed in the earliest releases but become progressively tighter as h increases. The interquartile range shrinks with each release, and the median gradually approaches zero, indicating that later releases tend to be closer to the final values. The same happens with the standard deviation and the average. The largest drop in dispersion occurs between the second and third releases, a feature consistent with the broad distribution observed for the third revision. A further though smaller adjustment takes place between the third and fourth releases. After this point, subsequent prediction errors cluster more tightly, with relatively small changes from one release to the next.

Regarding the distributions of revisions, early revisions are wide, asymmetric, and in some cases shifted away from zero, whereas later revisions are far more concentrated, suggesting that much of the updating is completed in the first few releases. The particularly dispersed nature of the second revision ($h = 3$) and the nontrivial distribution of the third ($h = 4$) hint that these releases carry a disproportionate share of the adjustment process. Both panels of Figure 2 convey a clear picture: informational gains from revisions are front-loaded, most of the noise is reduced within the first three to four releases, and subsequent releases mainly refine the estimates with comparatively small adjustments. These patterns provide a useful backdrop and motivation for the statistical analysis developed next.

Figure 2 shows that both prediction errors and revisions decrease in magnitude and converge toward zero as the release horizon h increases. Although the figure only displays distributions up to $h \leq 9$, in practice almost all series converge to their terminal release within a year of revisions. Accordingly, we set $H = 12$ as the maximum horizon in our empirical analysis.

4.3 Rationality tests

4.3.1 Revisions

Table 1 reports the estimated coefficients from the testing equations (2) to (5) for releases $h = 2, \ldots, 12$.[5] The unbiasedness test (2) indicates that early releases systematically understate the final value of GDP growth, while the bias largely disappears in later rounds, consistent with releases converging to the final value. For early revisions ($h \leq 5$), the estimates of α are positive and statistically significant, mostly at the 1 percent level, with values averaging above 0.05. For later revisions, the estimates become small and often insignificant (revisions $h = 7$ and $h = 9$ are exceptions), with estimates for $h \geq 10$ very close to zero.

On the other hand, the serial correlation test (3) provides little evidence of strong persistence in revisions over time. The few significant cases suggest some "carry-over" or correction effects, but not a systematic pattern. In particular, the estimates of ρ are generally small and insignificant, except for $h = 3$ and $h = 9$, where they are significant at most at the 5 percent level.

[5] Under rationality, prediction errors have a moving average structure and are therefore autocorrelated. For this reason, heteroskedasticity- and autocorrelation-consistent standard errors are used in the rationality tests below. This is a conservative approach in the case of revisions: under rationality, revisions are white noises and no adjustment for autocorrelation – which typically inflates the standard errors – is required. The conclusions of the rationality tests using only heteroskedasticity-consistent standard errors are virtually the same as those presented here. These results are available upon request.

By contrast, the cross-release correlation test (4) reveals stronger and more systematic patterns. In particular, the γ coefficients are significantly negative, implying that large revisions in one release tend to be partially undone in the subsequent release, which suggests smoothing across releases. This is especially evident for $h = 4$, where $\gamma_4 = -0.665$ and significant at the 1 percent level, and for $h = 7$, albeit at the weaker 10 percent level. The estimate for $h = 10$ is also large and negative ($\gamma_{10} = -0.272$), though not statistically significant.

Finally, the omnibus test (5), by combining bias with both types of autocorrelation, provides a unified perspective showing that different releases may fail rationality for different reasons: early releases exhibit systematic upward bias, middle releases display strong predictability from both previous releases and previous periods, and rationality is eventually restored as the coefficients in final releases shrink toward zero.

What is more revealing, however, is that the augmented equations allow us to analyze how the two forms of autocorrelation interact. In particular, the results point to notable transitions in the revision dynamics from $h = 3$ to $h = 4$, and again from $h = 9$ to $h = 10$. At $h = \{3, 9\}$, the estimates show $\rho < 0$ ($\rho_3 = -0.073$, $\rho_9 = -0.049$) and $\gamma = 0$, consistent with a correction mechanism in which revisions tend to reverse direction over time. The pattern then flips in $h = \{4, 10\}$ to $\rho > 0$ ($\rho_4 = 0.099$, $\rho_{10} = 0.195$) and $\gamma < 0$ ($\gamma_4 = -0.685$, $\gamma_{10} = -0.249$), implying that revisions become persistent across time while alternating across consecutive releases. The recurrence of such shifts indicates that structural updates periodically reset the mechanism governing revisions, a point we pursue in the next subsection.

4.3.2 Benchmarking

The systematic shifts described above are difficult to reconcile with purely efficient processing of new data, and may also reflect measurement error, evolving estimation procedures, or institutional constraints in data processing. One natural interpretation, which can be assessed directly with the data, is that the statistical agency engages in benchmarking, whereby monthly figures are adjusted to align with quarterly or annual targets. This practice induces characteristic dependence across revisions and prediction errors: negative release correlations ($\gamma < 0$) consistent with smoothing across releases, positive time-series correlations ($\rho > 0$) consistent with gradual corrections over time, and in some cases the joint appearance of both. Such dynamics point to a process in which abrupt corrections are avoided by partially reversing the previous vintage while still moving in the same direction across consecutive months, thereby keeping the path toward the benchmark smooth.

The evidence from Table 1 fits closely with this interpretation. In particular, in our sample, almost every three months (at $h = \{3, 6, 9\}$), monthly figures are released alongside quarterly totals, and the subsequent month (at $h = \{4, 7, 10\}$) typically marks the beginning of a new benchmarking exercise. The fact that such patterns emerge in consecutive revisions (e.g., from $h = 3$ to $h = 4$, and from $h = 9$ to $h = 10$) is consistent with cycles in which revisions first move strongly in one direction and are then partially corrected, while the overall trajectory continues to drift toward the benchmark.

These benchmarking cycles are not only statistical artifacts but also the result of major informational shocks. Quarterly releases provide the agency with more structured and comprehensive data, often derived from consolidated sectoral surveys, which affect multiple months simultaneously. Faced with this richer information, the agency follows a sequential revision strategy: at the beginning of a cycle, corrections are often abrupt, reflecting the need to quickly adjust for errors that had accumulated in earlier months; as the cycle unfolds, adjustments become smoother and more persistent, with revisions spread across several releases to avoid disrupting the credibility of published figures. In the final stage of the cycle, the agency makes any remaining corrections to ensure that estimates converge to the definitive GDP figure. This sequence—early strong corrections, mid-cycle gradual adjustments, and late convergence—points to an institutional learning process in which the agency balances the twin goals of accuracy and credibility when integrating new information.

To investigate this directly, we augment the omnibus regression with a dummy variable Q_t that equals one when the monthly release coincides with a quarterly release:

$$r_t^h = \alpha + \gamma\, r_t^{h-1} + \rho\, r_{t-1}^h + (\alpha_Q + \gamma_Q\, r_t^{h-1} + \rho_Q\, r_{t-1}^h) \times Q_t + \varepsilon_t\,. \tag{13}$$

The coefficients α_Q, γ_Q, and ρ_Q measure how the baseline parameters change in benchmark months. Table 2 shows that these interaction terms are large and highly significant at many horizons, and their inclusion reduces the magnitude of the baseline coefficients, suggesting that benchmarking is an important driver of apparent irrationality. In particular, the estimates of α_Q are significant and positive for $h = \{4, 6, 9\}$, indicating that revisions are systematically larger when quarterly figures are released. The estimates of ρ_Q are large, negative, and significant for $h = \{4, 5, 10\}$, showing that the correlation structure of revisions changes sharply in benchmark months, often reflecting stronger smoothing.

In conclusion, Table 2 provides direct empirical support for benchmarking as a systematic driver of non-rationality in revisions. Revisions are larger and more correlated precisely when quarterly figures are released, forecast errors are more tightly linked to revisions in these periods, and the correlation structure across releases shifts significantly. Together, these findings confirm that irrationality in revisions and errors is not random but clustered around benchmark months, consistent with the institutional procedures of the statistical agency.

4.3.3 Prediction errors

We now move to prediction errors, which provide a complementary perspective for assessing both the performance and the rationality of preliminary releases. Table 3 reports the estimated coefficients from the testing equations (6) to (9) for releases $h = 2, \ldots, 12$. The unbiasedness test (2) indicates that early releases underestimate the final value of GDP growth, while the bias gradually diminishes as more releases are published. For early revisions ($h \leq 5$), the estimates of α are positive, statistically significant at the 1 percent level, and decreasing with values ranging from above 0.286 to 0.076. For later revisions, the estimates become small and often insignificant (revisions $h = 7$ and $h = 9$ are exceptions), with estimates for $h \geq 10$ very close to zero.

The Mincer and Zarnowitz test (7) shows that early prediction errors are correlated with the size of the releases themselves, suggesting irrationality: larger prediction tend to be associated with larger errors. By later releases, this inefficiency fades, consistent with rationality only by the end of the revision cycle. The slope θ is positive and significant at the 5 percent level up to $h = 4$, and significant at the 10 percent level up to $h = 6$. The magnitude and significance of the intercept α is smaller than waht was found in the unbiasedness test. Recall that GDP growth has been steadily positive throughout the sample, so part of the effect previously captured by α is now attributed to y_t^h.

On the other hand, the cross-release test (8) indicate that in key releases, predictions errors are predictable from the revisions. In particular, when a large revision occurs, the previous forecast error tends to move in the opposite direction. This suggests that revisions are correcting systematic mismeasurement in earlier releases, and is captured through significantly negative γ coefficients at $h = \{3, 6, 10\}$.

Finally, the omnibus test of equation (9) confirms broad irrationality in prediction errors for early and middle releases, with evidence of bias, inefficiency, and systematic connection to revisions. The magnitude and significance of estimated coefficients gradually fade by the final releases.

It is worth mentioning that the evidence from both revisions (Table 1) and prediction errors (Table 3) is broadly consistent. However, the revision-based tests provide the clearest and most direct view: early releases are biased, middle releases exhibit predictable smoothing across releases, and rationality is only achieved in the final releases. The forecast error tests echo these findings but in a less direct way, reflecting their role as an alternative lens favored by much of the literature.

It is worth mentioning that the evidence from both revisions (Table 1) and prediction errors (Table 3) is broadly consistent. The revision-based tests provide a clear and direct view: early releases are biased, middle releases exhibit predictable smoothing across releases, and rationality is only achieved in the final releases. Prediction error tests echo these patterns as well, offering a complementary perspective that reinforces and extends the evidence from revisions.

4.4 Nowcasting

The evidence against rationality provides a strong motivation for the nowcasting exercise described in section 3.4. Table 4 reports the estimation results of the nowcasting equation (10). The autoregressive coefficient is statistically significant, with values around $\delta \simeq 0.3$, especially in the early releases. Importantly, the inclusion of autoregressive terms does not alter the qualitative conclusions regarding the lack of rationality in prediction errors. Almost all coefficients that appear significant in the omnibus tests of Table 3 remain significant in Table 4.

Furthermore, there is no evidence of first-order serial correlation in the residuals, as indicated by Breusch-Godfrey statistics (line "BG") remaining well below the 10% asymptotic critical value of 2.71. This finding means that the equation is dynamically well-specified, thereby capturing all relevant patterns for prediction.

At the bottom of the table, we present statistics comparing the (unadjusted) release y_t^h with the (adjusted) nowcast $\hat{y}_t$, as defined in equation (12). The first measure is the ratio of root mean squared errors (RMSE), which captures the prediction gains of the nowcast (numerator) over the official release (denominator). As expected, prediction gains decline with h, but for the first three releases they reach about 15% relative to the RMSE of the official release. The second measure is the Diebold and Mariano (1995) test, which evaluates whether the difference in predictive accuracy is statistically significant. The test statistic is the z-score of the loss (here, mean squared error) differential between the nowcast and the release. The negative sign of the statistic confirms that the nowcasts outperform the releases, with the differences being statistically significant for the initial three releases.

The last two statistics correspond to the encompassing test of Hendry and Clements (2004).

Specifically, these are the t-statistics for testing $H_0 : \psi = 0$ and $H_0 : \psi = 1$, where ψ is the weight that optimizes the combined prediction $\psi \hat{y}_t + (1-\psi) y_t^h$. The results show sustained rejections of $H_0 : \psi = 0$ and generalized non-rejection of $H_0 : \psi = 1$, indicating that ψ is statistically significant and close to one. In other words, the nowcast fully encompasses the official release, and combining the two predictions yields no additional gains.

These results confirm that the corrections introduced by the nowcasting procedure are not only statistically significant but also economically relevant, establishing nowcasting based solely on the information available from releases and revisions as a practical tool for improving the accuracy of early GDP growth estimates. Notably, simpler approaches can often rival or even outperform more sophisticated models (Banerjee and Marcellino, 2006). In our case, the nowcasting exercise achieves gains of up to 16.4% in relative accuracy (see $h = 3$ in Table 4), a result that is comparable to those reported by Strohsal and Wolf (2020) using a considerably more complex framework.

Figure 3 illustrates the performance of the nowcasting adjustment for selected representative releases. Each panel compares the official release trajectory of GDP growth (solid lines) with the corresponding adjusted nowcast (dashed lines). Panel (a) shows results for May 2006 and July 2012, while panel (b) reports the same comparison for August 2015 and July 2018. In both cases, the adjusted nowcasts move ahead of the vintage paths, converging more rapidly toward the final estimates. This illustrates how the nowcasting procedure accelerates the information content of early releases, improving predictive accuracy even when initial official releases are relatively noisy.

5 Concluding remarks

This paper documents and analyzes the revision process of monthly GDP growth in Peru. Using a newly constructed real-time dataset based on official publications, we evaluate the rationality of the statistical releases through a series of regression-based tests. The evidence reveals systematic deviations from rationality in early and intermediate releases: revisions are biased, predictable, and display significant vintage and time-series dependence. These patterns are particularly pronounced around the months in which quarterly figures are released, which points to benchmarking practices as a key source of revision dynamics.

Methodologically, our contribution is twofold. First, we propose a unified testing framework that serves as a diagnostic tool to uncover the different ways in which information is inefficiently processed. Second, and most importantly, we show that this same framework can be directly inverted to produce real-time corrections for preliminary estimates. Our adjustment procedure is simple, transparent, and yields statistically and economically significant gains in predictive accuracy, producing nowcasts that fully encompass the information in the original releases.

Beyond the specific case of Peru, our findings highlight the importance of documenting and studying revision processes in emerging economies, not just as technical artifacts but as reflections of institutional practices that can be turned into forecasting gains in real time. The lack of official vintage repositories should not prevent meaningful analysis: even in challenging data environments, it is possible to reconstruct revision histories and evaluate the performance of statistical agencies. We hope that our data collection strategy, rationality framework, and nowcasting method can serve as useful tools for future research in other countries facing similar data constraints.

References

Amir-Ahmadi, P., Matthes, C., and Wang, M.-C. (2017). Measurement errors and monetary policy: Then and now. *Journal of Economic Dynamics and Control*, 79:66–78.

Aoki, K. (2003). On the optimal monetary policy response to noisy indicators. *Journal of Monetary Economics*, 50:510–523.

Aruoba, S. B. (2008). Data revisions are not well behaved. *Journal of Money, Credit, and Banking*, 40:319–340.

Aruoba, S. B., Diebold, F. X., Nalewaik, J., Schorfheide, F., and Song, D. (2016). Improving gdp measurement: A measurement-error perspective. *Journal of Econometrics*, 191(2):384–397. Innovations in Measurement in Economics and Econometrics.

Bakhshi, H., Kapetanios, G., and Yates, T. (2005). Rational expectations and fixed-event forecasts: An application to UK inflation. *Empirical Economics*, 30:539–553.

Banerjee, A. and Marcellino, M. (2006). Are there any reliable leading indicators for us inflation and gdp growth? *International Journal of Forecasting*, 22(1):137–151.

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). Now-casting and the real-time data flow. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, chapter 4, pages 195–237. Elsevier.

Bernhard, S. (2016). A real-time GDP data set for Switzerland. Economic Studies 2016-09, Swiss National Bank.

Bernhardsen, T., Eitrheim, O., Jore, A. S., and Roisland, O. (2005). Real time data for Norway: Challenges for monetary policy. *North American Journal of Economics and Finance*, 16(3):333–49.

Bordignon, S. and Trivellato, U. (1989). The optimal use of provisional data in forecasting with dynamic models. *Journal of Business and Economic Statistics*, 7(2):275–286.

Castro, F. D., Perez, J. J., and Rodriguez-Vives, M. (2013). Fiscal data revisions in Europe. *Journal of Money, Credit, and Banking*, 45(6):1187–1209.

Cimadomo, J. (2016). Real-time data and fiscal policy analysis: A survey of the literature. *Journal of Economic Surveys*, 30(2):302–326.

Clements, M. P. (1997). Evaluating the rationality of fixed-event forecasts. *Journal of Forecasting*, 16(4):225–239.

Clements, M. P. (2022). Individual forecaster perceptions of the persistence of shocks to GDP. *Journal of Applied Econometrics*, 37(3):640–656.

Clements, M. P. and Harvey, D. I. (2010). Forecast encompassing tests and probability forecasts. *Journal of Applied Econometrics*, 25(6):1028–1062.

Conrad, W. and Corrado, C. (1979). Application of the Kalman filter to revisions in monthly retail sales estimates. *Journal of Economic Dynamics and Control*, 1:177–98.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. In Elliott, G., Granger, C., and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 1, chapter 17, pages 961–982. Elsevier.

Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of Economic Literature*, 49(1):72–100.

Croushore, D. and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, 105:111–130.

Croushore, D. and Stark, T. (2003). A real-time data set for macroeconomists: Does the data vintage matter? *Review of Economics and Statistics*, 85:605–617.

Cunningham, A., Eklund, J., Jeffery, C., Kapetanios, G., and Labhard, V. (2012). A state space approach to extracting the signal from uncertain data. *Journal of Business and Economic Statistics*, 30(2):173–180.

DeJong, P. (1987). Rational economic data revisions. *Journal of Business and Economic Statistics*, 5:539–548.

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

Ducoudré, B., Hubert, P., and Tabarly, G. (2020). The state-dependence of output revisions. *Economics Letters*, 192:109223.

Durbin, J. and Quenneville, B. (1997). Benchmarking by state space models. *International Statistical Review*, 65(1):23–48.

Egginton, D. M., Pick, A., and Vahey, S. P. (2002). 'Keep it real!': A real-time UK macro data set. *Economics Letters*, 77:15–20.

Faust, J., Rogers, J. H., and Wright, J. H. (2005). News and noise in G-7 GDP announcements. *Journal of Money, Credit, and Banking*, 37:403–419.

Fernandez, A. Z., Koenig, E. F., and Nikolsko-Rzhevskyy, A. (2011). A real-time historical database for the oecd. published online.

Garratt, A., Lee, K., Mise, E., and Shields, K. (2008). Real time representations of the output gap. *Review of Economics and Statistics*, 90:792–804.

Garratt, A. and Vahey, S. P. (2006). UK real-time macro data characteristics. *Economic Journal*, 116:F119–F135.

Giannone, D., Henry, J., Lalik, M., and Modugno, M. (2012). An area wide real time data base for the euro area. *The Review of Economics and Statistics*, 94(4):1000–1013.

Guerrero, V. M. (1993). Combining historical and preliminary information to obtain timely time series data. *International Journal of Forecasting*, 9(4):477–485.

Hecq, A., Jacobs, J. P., and Stamatogiannis, M. P. (2019). Testing for news and noise in non-stationary time series subject to multiple historical revisions. *Journal of Macroeconomics*, 60:396–407.

Hendry, D. F. and Clements, M. P. (2004). Pooling of forecasts. *Econometrics Journal*, 7(1):1–31.

Howrey, E. P. (1978). The use of preliminary data in econometric forecasting. *Review of Economics and Statistics*, 60:193–200.

INEI (2016). Methodology for calculating thr monthly index of national production. Technical report, National Institute of Statistics and Informatics - Technical Directorate of Economic Indicators, Lima.

INEI (2021). Metodología de cálculo del Índice mensual de la producción nacional. Technical report, Instituto Nacional de Estadística e Informática - Dirección Técnica de Indicadores Económicos.

Jacobs, J. P. and van Norden, S. (2011). Modeling data revisions: Measurement error and dynamics of 'true' values. *Journal of Econometrics*, 161(2):101–109.

Jacobs, J. P. and van Norden, S. (2016). Why are initial estimates of productivity growth so unreliable? *Journal of Macroeconomics*, 47:200–213.

Jacobs, J. P. A. M., Sarferaz, S., Sturm, J.-E., and van Norden, S. (2022). Can GDP measurement be further improved? data revision and reconciliation. *Journal of Business and Economic Statistics*, 40(1):423–431.

Kapetanios, G. and Yates, T. (2010). Estimating time variation in measurement error from data revisions: An application to backcasting and forecasting in dynamic models. *Journal of Applied Econometrics*, 25(5):869–893.

Kishor, N. K. (2011). Data revisions in India: Implications for monetary policy. *Journal of Asian Economics*, 22(2):164–173.

Kishor, N. K. and Koenig, E. F. (2012). VAR estimation and forecasting when data are subject to revision. *Journal of Business and Economic Statistics*, 30:181–190.

Kugler, P., Jordan, T. J., Lenz, C., and Savioz, M. R. (2005). GDP data revisions and forward-looking monetary policy in Switzerland. *North American Journal of Economics and Finance*, 16:351–372.

Lee, K. (2018). *Quarterly National Accounts Manual (2017 Edition)*. International Monetary Fund, Washington DC.

Mankiw, N. G., Runkle, D. E., and Shapiro, M. D. (1984). Are preliminary announcements of the money stock rational forecasts? *Journal of Monetary Economics*, 14(1):15–27.

Mankiw, N. G. and Shapiro, M. D. (1986). News or noise: An analysis of GNP revisions. *Survey of Current Business*, 66(May):20–25.

Mincer, J. A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. In Mincer, J. A., editor, *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, Studies in Business Cycles, chapter 1, pages 3–46. NBER and Columbia University Press.

Mork, K. A. (1987). Ain't behavin': Forecast errors and measurement errors in early GNP estimates. *Journal of Business and Economic Statistics*, 5(2):165–75.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Nordhaus, W. D. (1987). Forecasting efficiency: Concepts and applications. *Review of Economics and Statistics*, 69(4):667–674.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, 91(4):964–985.

Patterson, K. (1995). An integrated model of data measurement and data generation processes with an application to consumers' expenditure. *Economic Journal*, 105:54–76.

Patterson, K. (2003). Exploiting information in vintages of time-series data. *International Journal of Forecasting*, 19:177–197.

Patton, A. J. and Timmermann, A. (2012). Forecast rationality tests based on multi-horizon bounds. *Journal of Business and Economic Statistics*, 30(1):1–17.

Ramos, R. L. O., Robitaille, P. T., and de la Rocque Palis, R. (2004). News or noise? an analysis of Brazilian GDP announcements. International Finance Discussion Papers 776, Board of Governors of the Federal Reserve System (U.S.).

Sim, B., de Castro, M., and Pascua, M. (2009). An analysis of revisions to annual GDP estimates of six ADB regional members. ADB Economics Working Paper Series 184, Asian Development Bank (ADB).

Stark, T. (2011). Real GDP in annual revisions to the U.S. national accounts: 1966-2011. *Federal Reserve Bank of Philadelphia Research Rap.*

Stark, T. and Croushore, D. (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics*, 24(4):507–31.

Strohsal, T. and Wolf, E. (2020). Data revisions to German national accounts: Are initial releases good nowcasts? *International Journal of Forecasting*, 36(4):1252–1259.

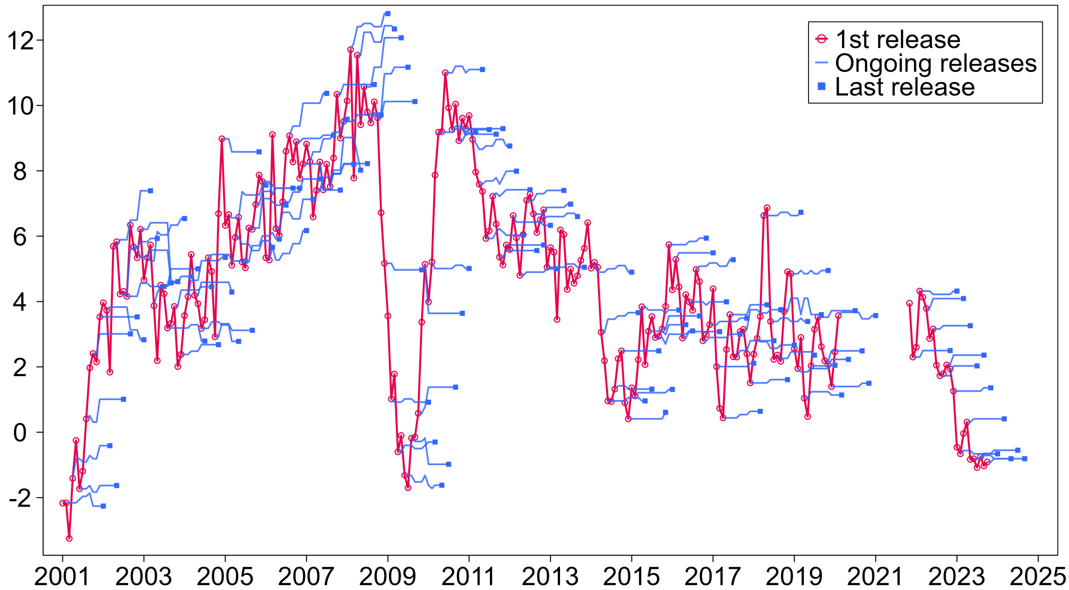
Weigand, R., Wanger, S., and Zapf, I. (2018). Factor structural time series models for official statistics with an application to hours worked in Germany. *Journal of Official Statistics*, 34(1):265–301.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838.

Yadav, A. (2021). Nowcasting Indian GDP in real time using factor bridge model and factor VAR. *Global Business Review*, 22(5):1289–1300.

Figures and Tables

Figure 1. *Initial releases and vintage paths of monthly GDP, 2001-2025*

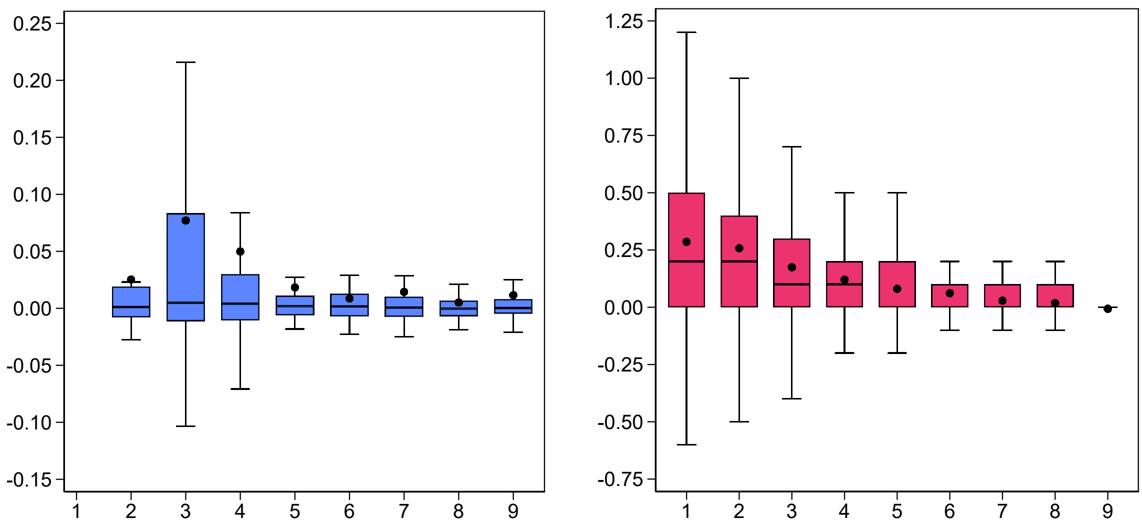


Notes: Initial releases of monthly GDP growth (red) and subsequent vintage paths (blue). Only even-month releases are shown to avoid clutter. Revisions in 2013 are excluded due to a base-year change, and the period 2020–2021 is omitted because of the undue volatility associated with COVID-19 lockdowns.

Figure 2. *Distribution of revisions and prediction errors across releases*

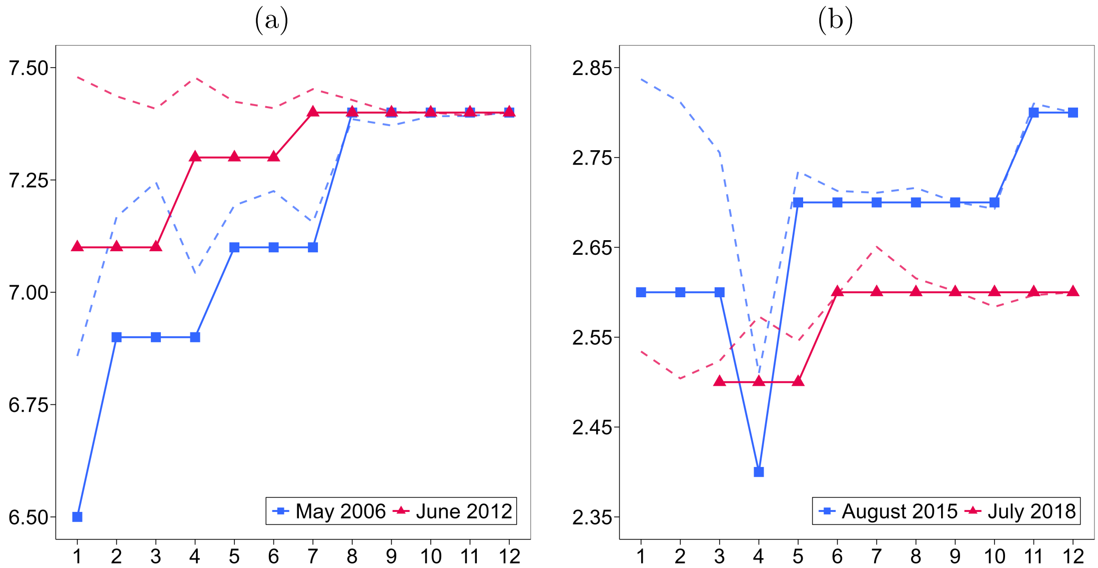
(a) Revisions

(b) Predicition errors



Notes: Boxplots of revisions and prediction errors across GDP releases. The averages are shown as filled circles.

Figure 3. *Nowcasting adjustment and vintage paths of monthly GDP (selected months)*



Notes: Comparison of official vintage paths (solid lines) and adjusted nowcasts (dashed lines) of monthly GDP growth for selected months.

Table 1. *Rationality tests based on revisions*

h	2	3	4	5	6	7	8	9	10	11	12
Unbiasedness											
α	0.028**	0.086***	0.055**	0.040***	0.019	0.031***	0.009	0.024***	−0.001	0.000	−0.006
	(0.012)	(0.022)	(0.022)	(0.010)	(0.013)	(0.012)	(0.007)	(0.009)	(0.013)	(0.000)	(0.009)
Serial correlation											
α	0.030**	0.094***	0.058**	0.041***	0.017	0.031**	0.009	0.025***	0.000	0.000	−0.005
	(0.013)	(0.023)	(0.024)	(0.011)	(0.013)	(0.013)	(0.007)	(0.010)	(0.010)	(0.000)	(0.009)
ρ	−0.038	−0.073**	−0.036	−0.030	0.001	−0.026	−0.018	−0.047***	0.220*	0.000	−0.001
	(0.034)	(0.032)	(0.025)	(0.023)	(0.022)	(0.022)	(0.017)	(0.017)	(0.118)	(0.000)	(0.024)
cross-release correlation											
α		0.088***	0.112***	0.042***	0.023	0.040***	0.010	0.024***	0.006	0.000	−0.006
		(0.023)	(0.020)	(0.010)	(0.015)	(0.013)	(0.007)	(0.009)	(0.010)	(0.000)	(0.008)
γ		−0.057	−0.665***	−0.026	−0.120	−0.445*	−0.023	−0.026	−0.283	−0.021	0.116
		(0.059)	(0.161)	(0.021)	(0.082)	(0.236)	(0.025)	(0.018)	(0.175)	(0.017)	(0.092)
Omnibus											
α		0.094***	0.111***	0.042***	0.021	0.037***	0.010	0.025***	0.006	0.000	−0.005
		(0.024)	(0.021)	(0.011)	(0.015)	(0.012)	(0.007)	(0.010)	(0.010)	(0.000)	(0.009)
γ		0.000	−0.689***	−0.025	−0.136	−0.457*	−0.022	0.003	−0.249**	−0.032	0.131
		(0.063)	(0.146)	(0.027)	(0.101)	(0.236)	(0.033)	(0.019)	(0.098)	(0.023)	(0.110)
ρ		−0.073**	0.101***	−0.006	0.044	0.036	−0.006	−0.049**	0.196***	0.033	−0.031
		(0.036)	(0.033)	(0.036)	(0.047)	(0.044)	(0.033)	(0.020)	(0.058)	(0.025)	(0.050)
N	239	239	239	239	239	239	239	239	239	239	239

Notes: Least squares estimation of equations (2), (3), (4), and (5). Heteroskedasticity- and autocorrelation-consistent standard errors (with 6 lags) in parentheses (cf. Newey and West, 1987). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. N is the number of observations.

Table 2. *Benchmarking*

h	3	4	5	6	7	8	9	10	11	12
α	0.061***	0.053**	0.020	−0.004	0.019	0.000	0.004	0.006	−0.001	−0.007
	(0.020)	(0.021)	(0.013)	(0.021)	(0.013)	(0.010)	(0.006)	(0.010)	(0.007)	(0.012)
γ	0.073	−0.045	−0.007	−0.194	−0.583**	0.003	−0.006	−0.025	−0.038	0.257
	(0.068)	(0.057)	(0.012)	(0.194)	(0.264)	(0.037)	(0.022)	(0.020)	(0.040)	(0.218)
ρ	−0.108*	0.001	0.009	0.059	0.066	−0.018	−0.050	0.012	−0.002	−0.030
	(0.060)	(0.014)	(0.037)	(0.070)	(0.073)	(0.049)	(0.036)	(0.008)	(0.023)	(0.055)
α_Q	0.081	0.112***	0.056*	0.060**	0.031	0.024	0.048**	−0.004	0.001	0.005
	(0.055)	(0.038)	(0.030)	(0.026)	(0.020)	(0.016)	(0.023)	(0.025)	(0.014)	(0.020)
γ_Q	0.038	0.061	0.055	−0.151*	−0.188**	0.058	−0.022	0.260***	0.079	0.081
	(0.072)	(0.078)	(0.064)	(0.083)	(0.086)	(0.067)	(0.055)	(0.086)	(0.070)	(0.076)
ρ_Q	−0.214	−0.774***	−0.149***	0.135	0.509*	−0.124*	0.004	−0.219**	−0.009	−0.269
	(0.151)	(0.091)	(0.040)	(0.192)	(0.263)	(0.073)	(0.039)	(0.102)	(0.051)	(0.220)
N	239	239	239	239	239	239	239	239	239	239

Notes: Least squares estimation of equation (13). Heteroskedasticity- and autocorrelation-consistent standard errors (with 6 lags) in parentheses (cf. Newey and West, 1987). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. N is the number of observations.

Table 3. *Rationality tests based on prediction errors*

h	1	2	3	4	5	6	7	8	9	10	11
Unbiasedness											
α	0.286***	0.257***	0.172***	0.117***	0.076***	0.058**	0.026	0.017	−0.007	−0.006	−0.006
	(0.054)	(0.048)	(0.044)	(0.032)	(0.026)	(0.027)	(0.023)	(0.020)	(0.021)	(0.012)	(0.009)
Mincer-Zarnowitz											
α	0.131**	0.097**	0.058	0.028	0.012	−0.002	−0.013	−0.001	−0.005	−0.014	−0.005
	(0.051)	(0.047)	(0.038)	(0.030)	(0.028)	(0.023)	(0.020)	(0.018)	(0.016)	(0.013)	(0.010)
θ	0.036**	0.037**	0.026**	0.020**	0.014*	0.013*	0.009	0.004	0.000	0.002	0.000
	(0.017)	(0.016)	(0.013)	(0.009)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)	(0.004)	(0.003)
cross-release correlation (encompassing)											
α		0.257***	0.221***	0.116***	0.075***	0.066**	0.033	0.018	−0.002	−0.006	−0.006
		(0.048)	(0.043)	(0.031)	(0.027)	(0.027)	(0.024)	(0.020)	(0.019)	(0.012)	(0.008)
γ		0.032	−0.579***	0.011	0.040	−0.433*	−0.191	−0.060	−0.203	−0.047**	0.116
		(0.167)	(0.217)	(0.040)	(0.098)	(0.261)	(0.120)	(0.112)	(0.163)	(0.018)	(0.092)
Omnibus											
α		0.096**	0.058	0.024	0.006	−0.002	−0.011	−0.002	−0.002	−0.012	−0.002
		(0.047)	(0.044)	(0.032)	(0.028)	(0.026)	(0.022)	(0.019)	(0.017)	(0.014)	(0.009)
θ		0.038**	0.035**	0.023**	0.014*	0.016**	0.011*	0.005	0.001	0.002	−0.001
		(0.016)	(0.014)	(0.010)	(0.008)	(0.008)	(0.007)	(0.006)	(0.005)	(0.004)	(0.003)
γ		0.003	−0.617***	−0.027	0.015	−0.449*	−0.222*	−0.081	−0.206	−0.028	0.118
		(0.182)	(0.209)	(0.033)	(0.099)	(0.248)	(0.125)	(0.115)	(0.162)	(0.019)	(0.093)
ρ		−0.059	0.134**	−0.153***	0.191	−0.121	−0.228*	−0.061	−0.069	−0.093***	−0.002
		(0.140)	(0.060)	(0.044)	(0.208)	(0.215)	(0.134)	(0.128)	(0.105)	(0.025)	(0.010)
N	239	239	239	239	239	239	239	239	239	239	239

Notes: Least squares estimation of equations (6), (7), (8), and (9). Heteroskedasticity- and autocorrelation-consistent standard errors (with 6 lags) in parentheses (cf. Newey and West, 1987). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. N is the number of observations.

Table 4. *Nowcasting equation and evaluation*

h	1	2	3	4	5	6	7	8	9	10	11
α	0.080**	0.059	0.033	0.016	0.004	−0.003	−0.008	−0.004	−0.001	−0.010	−0.001
	(0.040)	(0.037)	(0.034)	(0.028)	(0.027)	(0.024)	(0.021)	(0.020)	(0.018)	(0.014)	(0.010)
θ	0.024**	0.026**	0.026**	0.019**	0.012	0.013*	0.010*	0.005	0.000	0.001	−0.001
	(0.011)	(0.010)	(0.010)	(0.008)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)	(0.003)
γ		−0.217	−0.654***	−0.100	−0.044	−0.440*	−0.287*	−0.180	−0.154	−0.085	0.134
		(0.184)	(0.188)	(0.071)	(0.113)	(0.252)	(0.169)	(0.176)	(0.103)	(0.053)	(0.115)
ρ			0.316***	−0.154***	0.188	−0.042	−0.180	−0.051	−0.006	−0.073***	0.002
			(0.084)	(0.055)	(0.206)	(0.209)	(0.140)	(0.138)	(0.140)	(0.028)	(0.014)
δ	0.365***	0.354***	0.312***	0.242**	0.134	0.179*	0.246*	0.207	0.296**	0.165	−0.032
	(0.091)	(0.098)	(0.098)	(0.123)	(0.097)	(0.105)	(0.132)	(0.135)	(0.119)	(0.141)	(0.053)
BG	1.56	1.32	1.08	0.44	0.06	0.32	0.68	0.47	0.98	0.02	0.01
RMSE ratio	85.8	86.4	84.6	93.6	96.4	93.8	97.3	99.8	99.4	99.6	99.7
DM	−2.36***	−2.25**	−1.88**	−1.54*	−1.48*	−1.04	−1.23	−0.40	−0.86	−1.14	−0.59
ET $\psi = 0$	2.16**	2.15**	4.13***	3.20***	2.08**	2.49**	2.60***	1.09	1.64*	4.55***	1.27
ET $\psi = 1$	−0.31	0.12	0.11	0.32	0.04	−0.06	0.29	0.52	−0.46	1.59	0.17
N	239	239	239	239	239	239	239	239	239	239	239

Notes: Least squares estimation of equations (10). Heteroskedasticity consistent standard errors in parentheses (cf. White, 1980). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. N is the number of observations and BG is the Breusch-Godfrey χ^2_1-statistic for zero first-order autocorrelation in the residuals. DM is the z-score for the Diebold and Mariano (1995) test. ET denotes the encompassing tests with null hypothesis $H_0 : \psi = 0$ (release y^h_t encompasses nowcast $\hat{y}_t$) and $H_0 : \psi = 1$ ($\hat{y}_t$ encompasses y^h_t).

Online Supplement (Not intended for publication)

Nowcasting GDP using data revisions in an emerging economy

January 19, 2026

This supplemental document provides detailed information on the data sources, collection procedures, processing steps, and terminology used to construct the first real-time dataset (RTD) of Peruvian GDP growth. It focuses on the methodological workflow, quality-control standards, and replication resources that underpin the empirical analysis.

In Peru, the National Institute of Statistics and Informatics (INEI) produces and disseminates GDP statistics, while the Central Reserve Bank of Peru (BCRP) republishes these figures in its Weekly Report (WR). Although INEI is the primary source, the BCRP provides a clearer and more consistent record of releases, reflecting legal obligations under article 84 of the Peruvian Constitution and articles 2 and 74 of the BCRP's Organic Law, which require the periodic publication of key national macroeconomic statistics. Our data collection strategy draws on both sources to construct the RTD of Peruvian GDP growth.

S1 Revision Calendar

A formal publication calendar governs the initial release of GDP growth figures, which are published with a two-month lag relative to the reference period across all frequencies. While this delay aligns with international practice, it underscores the need to account for data revisions when using GDP for real-time policy analysis. In contrast, no official revision calendar is maintained. This distinction is important to highlight: the timing of first releases is regulated and predictable, whereas subsequent revisions follow no explicit or public schedule.

For the construction of a RTD, however, a revision calendar is indispensable. Without a consistent schedule of updates, it becomes impossible to define vintages and track revisions systematically. A regular calendar ensures comparability across time, while irregular release dates complicate measurement, since the horizon between releases varies.

To address this limitation, we reconstructed an implicit revision calendar using the BCRP's WR. Each WR reproduces INEI's figures and specifies the date of the update. We then harmonized this information to a monthly frequency—consistent with the "horizon" definition in the main paper—based on two criteria:

1. Chief Resolution No. 316-2003-INEI mandates that sectoral offices update their data at least quarterly (e.g., March, June, September, December). These offices also produce monthly statistical information for the Monthly Index of National Production (IMPN), suggeting that revisions are made monthly.
2. In the WR, revisions are effectively made on a month-to-month basis rather than week-to-week. Our reconstructed calendar confirms that the WR report updated GDP figures at least once per month.

S2 Raw Data Sources

Once the monthly revision calendar was established, we compiled the last WR of each month. Each monthly WR defines a *vintage* of GDP growth rates, which together form the basis of the RTD.

The WR republishes GDP growth rates by sector as released by INEI. These sectoral growth rates are reported year-on-year and presented in a well-organized table whose structure has changed little over the past three decades, although the table numbering has varied across issues. Two key tables were systematically extracted: one reporting monthly growth rates, and another reporting quarterly and annual growth rates

Historically, until the early 2000s, the WR was titled the "Weekly Bulletin", and no digital versions of either the Bulletin or the Report exist prior to 2013. Consequently, for the period spanning 2000–2012, the macroeconomic tables are only accessible in hardcover volumes available at the BCRP's Renzo Rossini Library, requiring in-person consultation. These volumes cannot be removed from the library, but they may be photocopied or photographed. This process is not always straightforward, as the historical WRs are often bound in large, heavy volumes.

From 2013 to the present, however, WRs have been freely accessible on the BCRP's website in PDF format, which has greatly facilitated data collection. The data are also presented in spreadsheets and in the BCRP data database, although only in their revised form.

S2.1 Scanned Data (pre–2013)

For the pre-2013 period, the WRs were scanned into PDF format, producing non-structured hardcopies. Consequently, the extraction process was far from trivial and followed three main steps: (i) Computer-based extraction, (ii) manual correction and verification, and (iii) construction of parsed structured tables feeding into automated vintage-building codes.

The scanned sources posed significant challenges:

- Warped or distorted text and misaligned table lines, often caused by hardcover bindings during scanning.
- Shadowing, opacity, or blur reducing legibility.
- Fragmented or incomplete tables.
- Highlighted entries interfering with optical recognition accuracy.
- Inconsistent font sizes across tables within the same report.
- Residual marks obscuring decimal points or digits.
- Ink and other physical spots disrupting recognition.

Figures S1 and S2 show examples of scanned tables exhibiting some of these difficulties.

Given these limitations, data extraction was particularly demanding and did not provide definite results. We employed Optical Character Recognition (OCR) using the open-source `Tesseract` engine in `Python`. Originally developed by Hewlett-Packard and now maintained by Google, `Tesseract` is widely used for digitizing printed material, including statistical reports.

Figure S1. *Table 70, WR 2 (January 2007): monthly growth rates*

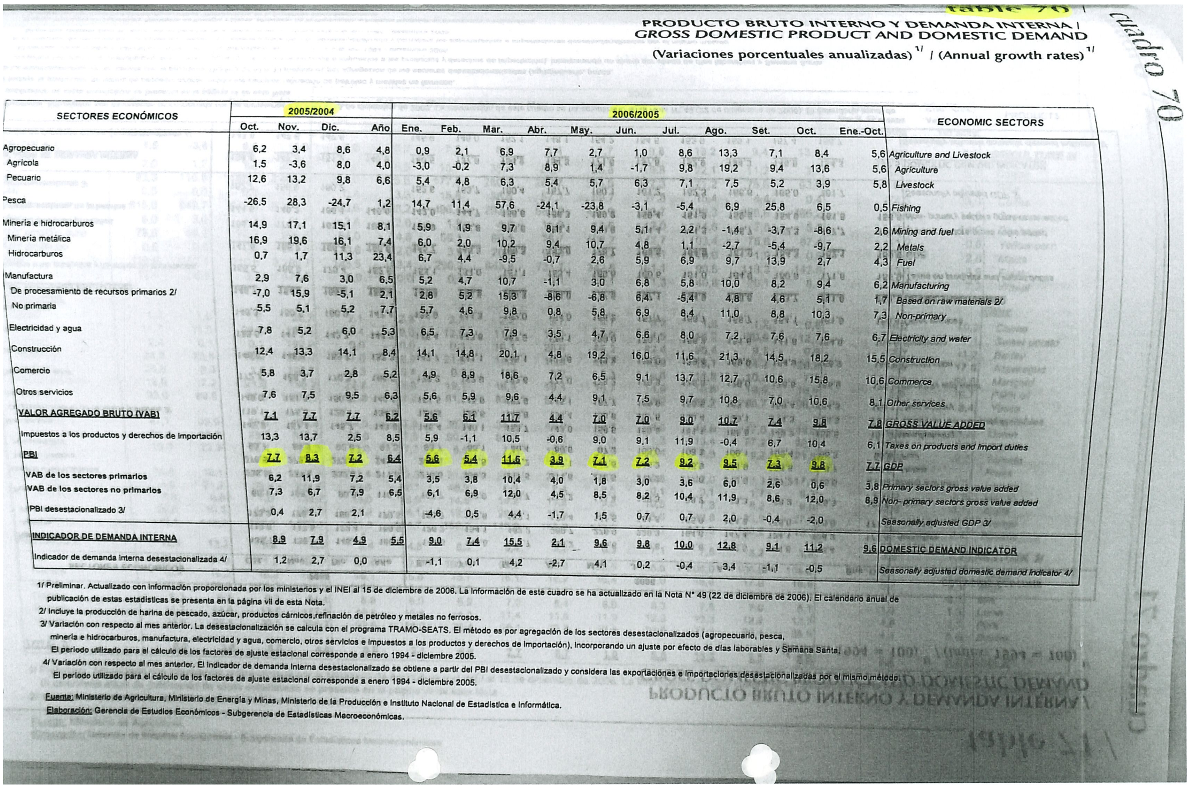


table 70

cuadro 70

PRODUCTO BRUTO INTERNO Y DEMANDA INTERNA /
GROSS DOMESTIC PRODUCT AND DOMESTIC DEMAND

(Variaciones porcentuales anualizadas) 1/ / (Annual growth rates) 1/

SECTORES ECONÓMICOS	2005/2004				2006/2005											ECONOMIC SECTORS
	Oct.	Nov.	Dic.	Año	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Set.	Oct.	Ene.-Oct.	
Agropecuario	6,2	3,4	8,6	4,8	0,9	2,1	6,9	7,7	2,7	1,0	8,6	13,3	7,1	8,4	5,6	*Agriculture and Livestock*
Agrícola	1,5	-3,6	8,0	4,0	-3,0	-0,2	7,3	8,9	1,4	-1,7	9,8	19,2	9,4	13,6	5,6	*Agriculture*
Pecuario	12,6	13,2	9,8	6,6	5,4	4,8	6,3	5,4	5,7	6,3	7,1	7,5	5,2	3,9	5,8	*Livestock*
Pesca	-26,5	28,3	-24,7	1,2	14,7	11,4	57,6	-24,1	-23,8	-3,1	-5,4	6,9	25,8	6,5	0,5	*Fishing*
Minería e hidrocarburos	14,9	17,1	15,1	8,1	5,9	1,9	9,7	8,1	9,4	5,1	2,2	-1,4	-3,7	-8,6	2,6	*Mining and fuel*
Minería metálica	16,9	19,6	16,1	7,4	6,0	2,0	10,2	9,4	10,7	4,8	1,1	-2,7	-5,4	-9,7	2,2	*Metals*
Hidrocarburos	0,7	1,7	11,3	23,4	6,7	4,4	-9,5	-0,7	2,6	5,9	6,9	9,7	13,9	2,7	4,3	*Fuel*
Manufactura	2,9	7,6	3,0	6,5	5,2	4,7	10,7	-1,1	3,0	6,8	5,8	10,0	8,2	9,4	6,2	*Manufacturing*
De procesamiento de recursos primarios 2/	-7,0	15,9	-5,1	2,1	2,8	5,2	15,3	-8,6	-6,8	6,4	-5,4	4,8	4,6	5,1	1,7	*Based on raw materials 2/*
No primaria	5,5	5,1	5,2	7,7	5,7	4,6	9,8	0,8	5,8	6,9	8,4	11,0	8,8	10,3	7,3	*Non-primary*
Electricidad y agua	7,8	5,2	6,0	5,3	6,5	7,3	7,9	3,5	4,7	6,6	8,0	7,2	7,6	7,6	6,7	*Electricity and water*
Construcción	12,4	13,3	14,1	8,4	14,1	14,8	20,1	4,8	19,2	16,0	11,6	21,3	14,5	18,2	15,5	*Construction*
Comercio	5,8	3,7	2,8	5,2	4,9	8,9	18,6	7,2	6,5	9,1	13,7	12,7	10,6	15,8	10,6	*Commerce*
Otros servicios	7,6	7,5	9,5	6,3	5,6	5,9	9,6	4,4	9,1	7,5	9,7	10,8	7,0	10,6	8,1	*Other services*
VALOR AGREGADO BRUTO (VAB)	**7,1**	**7,7**	**7,7**	**6,2**	**5,6**	**6,1**	**11,7**	**4,4**	**7,0**	**7,0**	**9,0**	**10,7**	**7,4**	**9,8**	**7,8**	***GROSS VALUE ADDED***
Impuestos a los productos y derechos de importación	13,3	13,7	2,5	8,5	5,9	-1,1	10,5	-0,6	9,0	9,1	11,9	-0,4	6,7	10,4	6,1	*Taxes on products and import duties*
PBI	**7,7**	**8,3**	**7,2**	**6,4**	**5,6**	**5,4**	**11,6**	**3,9**	**7,1**	**7,2**	**9,2**	**9,5**	**7,3**	**9,8**	**7,7**	***GDP***
VAB de los sectores primarios	6,2	11,9	7,2	5,4	3,5	3,8	10,4	4,0	1,8	3,0	3,6	6,0	2,6	0,6	3,8	*Primary sectors gross value added*
VAB de los sectores no primarios	7,3	6,7	7,9	6,5	6,1	6,9	12,0	4,5	8,5	8,2	10,4	11,9	8,6	12,0	8,9	*Non- primary sectors gross value added*
PBI desestacionalizado 3/	0,4	2,7	2,1		-4,6	0,5	4,4	-1,7	1,5	0,7	0,7	2,0	-0,4	-2,0		*Seasonally adjusted GDP 3/*
INDICADOR DE DEMANDA INTERNA	**8,9**	**7,9**	**4,9**	**5,5**	**9,0**	**7,4**	**15,5**	**2,1**	**9,6**	**9,8**	**10,0**	**12,8**	**9,1**	**11,2**	**9,6**	**DOMESTIC DEMAND INDICATOR**
Indicador de demanda interna desestacionalizada 4/	1,2	2,7	0,0		-1,1	0,1	4,2	-2,7	4,1	0,2	-0,4	3,4	-1,1	-0,5		*Seasonally adjusted domestic demand indicator 4/*

1/ Preliminar. Actualizado con información proporcionada por los ministerios y el INEI al 15 de diciembre de 2006. La información de este cuadro se ha actualizado en la Nota N° 49 (22 de diciembre de 2006). El calendario anual de publicación de estas estadísticas se presenta en la página vii de esta Nota.

2/ Incluye la producción de harina de pescado, azúcar, productos cárnicos,refinación de petróleo y metales no ferrosos.

3/ Variación con respecto al mes anterior. La desestacionalización se calcula con el programa TRAMO-SEATS. El método es por agregación de los sectores desestacionalizados (agropecuario, pesca, minería e hidrocarburos, manufactura, electricidad y agua, comercio, otros servicios e impuestos a los productos y derechos de importación), incorporando un ajuste por efecto de días laborables y Semana Santa.
El periodo utilizado para el cálculo de los factores de ajuste estacional corresponde a enero 1994 - diciembre 2005.

4/ Variación con respecto al mes anterior. El indicador de demanda interna desestacionalizado se obtiene a partir del PBI desestacionalizado y considera las exportaciones e importaciones desestacionalizadas por el mismo método.
El periodo utilizado para el cálculo de los factores de ajuste estacional corresponde a enero 1994 - diciembre 2005.

Fuente: Ministerio de Agricultura, Ministerio de Energía y Minas, Ministerio de la Producción e Instituto Nacional de Estadística e Informática.

Elaboración: Gerencia de Estudios Económicos - Subgerencia de Estadísticas Macroeconómicas.

Notes: The low resolution, uneven lighting, and shadowing from the page fold in this figure illustrate common difficulties for OCR processing. Distortions in the horizontal and vertical alignment of cells, faint characters, and overlapping text reduce recognition accuracy. Highlighted values and residual marks (e.g., spots or smudges) can obscure digits or decimal points, leading to frequent misclassifications. In this specific case, blurred numbers in the central columns and shading near the binding make it particularly challenging to extract reliable machine-readable data without extensive manual correction.

The OCR workflow consisted of:

1. Converting pages into PNG format.
2. Transforming visual text into machine-readable formats, with extensive pre-processing to improve accuracy:
   - Grayscale conversion and binarization to separate text from background.
   - Noise removal (e.g., speckles, smudges) using median filtering or adaptive thresholding.
   - Skew correction (deskewing) to align rotated text.
   - Image scaling to ensure at least 300 DPI resolution.
   - Contrast enhancement (e.g., via CLAHE) to make characters more prominent.
   - Thinning and skeletonization to separate thick or touching characters.
   - Border removal to avoid false detections.
3. Extracting tabular text into editable formats.

Although OCR provided an initial transcription, frequent misclassifications occurred (e.g., reading `9.4` as `5.4`), making manual verification indispensable.

Figure S2. *Table 88, WR 2 (January 2007): quarterly and annual growth rates*

table 88

PRODUCTO BRUTO INTERNO / *GROSS DOMESTIC PRODUCT*

(Variaciones porcentuales) [1/] / *(Percentages changes)* [1/]

SECTORES ECONÓMICOS	2004					2005					2006			ECONOMIC SECTORS
	I	II	III	IV	AÑO	I	II	III	IV	AÑO	I	II	III	
Agropecuario	4,6	-1,6	1,2	4,8	1,7	2,2	5,8	5,0	6,1	4,8	3,4	3,8	9,7	*Agriculture and Livestock*
Agrícola	1,2	-8,1	-3,2	2,3	-3,2	0,9	6,6	4,5	1,8	4,0	1,7	2,8	12,5	*Agriculture*
Pecuario	2,5	1,0	2,0	2,7	2,0	4,6	4,1	6,0	11,8	6,6	5,5	5,8	6,6	*Livestock*
Pesca	14,0	36,3	46,6	38,1	33,9	13,6	2,1	-6,4	-2,6	1,2	27,6	-17,7	5,8	*Fishing*
Minería e Hidrocarburos	12,3	2,8	1,5	4,8	5,2	1,0	4,8	10,9	15,7	8,1	5,9	7,5	-1,0	*Mining and fuel*
Minería metálica	13.2	2,9	0,8	4,4	5,2	-1,5	2,7	10,6	17,5	7,4	6,1	8,3	-2,4	*Metals*
Hidrocarburos	-5,4	-4,3	12,7	25,5	7,1	33,9	37,7	23,6	4,4	23,4	0,4	2,7	10,0	*Fuel*
Manufactura	5,4	6,2	6,9	10,9	7,4	7,3	8,0	6,2	4,5	6,5	6,9	2,9	8,0	*Manufacturing*
De procesamiento de recursos primarios	4,2	9,9	6,2	8,2	7,3	2,4	3,2	0,9	2,0	2,1	7,8	-3,2	1,0	*Based on raw materials 2/*
No primaria	5,7	5,1	7,1	11,7	7,4	8,6	9,5	7,7	5,3	7,7	6,7	4,5	9,4	*Non-primary*
Electricidad y agua	4,7	4,4	4,1	5,0	4,6	3,7	5,6	5,7	6,3	5,3	7,2	4,9	7,6	*Electricity and water*
Construcción	7,3	4,1	3,4	4,1	4,7	2,2	6,9	10,6	13,3	8,4	16,3	13,2	16,0	*Construction*
Comercio	2,9	4,0	5,4	11,4	5,8	6,3	5,9	4,4	4,1	5,2	10,9	7,6	12,3	*Commerce*
Otros servicios	3,3	3,2	4,8	6,4	4,4	6,2	4,9	5,8	8,3	6,3	7,1	7,0	9,2	*Other services*
VALOR AGREGADO BRUTO (VAB)	**4,7**	**3,4**	**4,7**	**7,7**	**5,1**	**5,4**	**5,8**	**6,2**	**7,5**	**6,2**	**7,8**	**6,1**	**9,0**	***GROSS VALUE ADDED***
Impuestos a los productos y derechos de importación	4,1	12,7	5,6	3,6	6,4	10,1	7,2	7,3	9,4	8,5	5,2	5,8	5,8	*Taxes on products and import duties*
PBI	**4,6**	**4,2**	**4,8**	**7,2**	**5,2**	**5,9**	**5,9**	**6,3**	**7,7**	**6,4**	**7,6**	**6,1**	**8,7**	***GDP***
VAB de los sectores primarios	7,4	2,5	2,9	6,4	4,6	2,0	4,9	6,1	8,5	5,4	5,7	2,9	4,1	*Primary sectors gross aggregated value*
VAB de los sectores no primarios	3,9	3,7	5,2	8,0	5,2	6,3	6,0	6,2	7,3	6,5	8,4	7,1	10,3	*Non- primary sectors gross aggregated value*

1/ Preliminar. Actualizado con información proporcionada por el INEI al 15 de noviembre de 2006. La información de este cuadro se ha actualizado en la Nota N° 45 (24 de noviembre de 2006).
El calendario anual de publicación de estas estadísticas se presenta en la página vii de esta Nota.

Fuente: Instituto Nacional de Estadística e Informática y Ministerios de Agricultura, Energía y Minas y de la Producción.
Elaboración: Gerencia de Estudios Económicos - Subgerencia de Estadísticas Macroeconómicas.

Notes: Another example of a problematic scanned WR table. OCR processing is hindered by uneven contrast, faint print, shading near the binding, and inconsistent alignment. Highlighted values, small font sizes, and residual marks further obscure digits and decimal points, making reliable digitization difficult without extensive pre-processing and manual checks.

Once raw tables were extracted and put in readable CSV format, we parsed and processed the data through the pipeline:

1. Data cleaning and validation using more than 70 specialized transformation functions.
2. Vintage-format conversion with target-event column structure.
3. Concatenation across vintages to form a complete RTD.
4. Storage in both CSV and Parquet formats for accessibility.

The combination of automated routines and manual oversight was essential for ensuring the reliability of the early vintages.

S2.2 Digital Data (post-2013)

Since 2013, WRs have been published online as digital PDFs (semi-structured data), greatly facilitating automated extraction. We developed `Python` routines to handle heterogeneous table structures across reports. Although digital sources were more consistent than scanned volumes, irregular formatting still required flexible parsing scripts. The primary extraction library was `tabula`, selected for its robustness in dealing with complex table layouts.

Figure S3. *Table 62, WR 28 (July 2013): monthly growth rates*

table 62 / cuadro 62

PRODUCTO BRUTO INTERNO Y DEMANDA INTERNA /
GROSS DOMESTIC PRODUCT AND DOMESTIC DEMAND

(Variaciones porcentuales anualizadas) 1/ / *(Annual growth rates)* 1/

SECTORES ECONÓMICOS										2013						ECONOMIC SECTORS
	May.	Jun.	Jul.	Ago.	Set.	Oct.	Nov.	Dic.	Año	Ene.	Feb.	Mar.	Abr.	May.	Ene. - May.	
Agropecuario	10,8	11,3	3,2	4,4	4,8	3,2	3,9	7,4	5,1	7,5	6,5	5,8	3,1	-0,7	3,7	*Agriculture and Livestock*
Agrícola	14,0	15,5	0,8	2,5	3,8	3,8	4,4	11,0	5,2	10,7	9,3	8,8	3,0	-2,0	4,2	*Agriculture*
Pecuario	3,8	3,6	6,9	6,8	5,8	2,5	3,3	3,4	4,9	3,9	3,5	2,0	3,5	2,4	3,1	*Livestock*
Pesca	5,2	20,4	0,6	-12,0	18,9	15,4	-15,8	-47,2	-11,9	6,5	1,3	-20,4	-9,7	-19,8	-9,8	*Fishing*
Minería e hidrocarburos	1,3	4,9	4,5	0,6	5,1	-3,2	-1,4	-1,6	2,2	-4,4	-1,7	3,4	7,8	5,9	2,3	*Mining and fuel*
Minería metálica	1,8	4,7	4,7	0,2	4,7	-5,2	0,4	-2,6	2,1	-8,6	-3,2	3,3	6,9	4,8	0,7	*Metals*
Hidrocarburos	-0,6	6,1	3,6	2,7	7,1	5,3	-9,1	3,1	2,3	16,3	4,7	3,9	12,1	10,7	9,4	*Fuel*
Manufactura	2,9	1,5	5,0	4,4	1,6	4,5	4,2	-2,0	1,3	3,1	0,2	-3,6	4,4	1,0	1,0	*Manufacturing*
De recursos primarios	-9,0	-2,0	-1,1	-5,1	-4,4	6,9	-3,3	-20,5	-6,5	-2,4	0,5	-5,6	-9,6	2,4	-2,9	*Based on raw materials*
No primaria	5,4	2,2	6,1	6,1	2,5	4,2	5,4	2,1	2,8	4,1	0,1	-3,3	6,9	0,7	1,6	*Non-primary*
Electricidad y agua	5,3	4,7	5,7	5,3	4,1	3,7	5,6	4,3	5,2	6,3	4,0	4,1	6,4	5,7	5,3	*Electricity and water*
Construcción	15,8	19,7	21,4	17,4	19,2	16,1	16,8	5,3	15,2	18,4	14,3	3,8	26,5	10,5	14,5	*Construction*
Comercio	6,5	6,7	6,9	6,5	5,3	5,9	6,3	7,0	6,7	5,5	5,6	4,0	7,5	6,6	5,9	*Commerce*
Otros servicios 2/	7,5	7,6	7,9	6,7	6,5	8,0	7,6	6,0	7,3	7,0	5,7	4,3	7,2	6,1	6,1	*Other services 2/*
PBI	**7,1**	**7,4**	**7,5**	**6,5**	**6,3**	**6,8**	**6,8**	**4,3**	**6,3**	**6,4**	**5,1**	**3,0**	**7,7**	**5,0**	**5,4**	***GDP***
Sectores primarios	4,9	7,6	2,8	1,0	3,6	1,7	0,4	-3,5	1,7	1,5	2,6	2,5	2,3	0,9	1,9	*Primary sectors*
Sectores no primarios	7,6	7,4	8,4	7,4	6,7	7,6	7,8	5,6	7,1	7,2	5,5	3,1	8,7	5,8	6,1	*Non- primary sectors*
PBI desestacionalizado 3/	**1,3**	**0,5**	**0,6**	**0,2**	**0,2**	**0,7**	**0,4**	**-0,3**		**1,2**	**0,5**	**0,1**	**0,3**	**0,6**		*Seasonally adjusted GDP 3/*
INDICADOR DE DEMANDA INTERNA	**9,7**	**7,5**	**12,8**	**9,7**	**6,7**	**9,7**	**7,2**	**6,0**	**7,4**	**11,6**	**9,3**	**4,2**	**8,9**	**6,0**	**7,9**	**DOMESTIC DEMAND INDICATOR**
Indicador de demanda interna desestacionalizada 3/	**1,5**	**1,0**	**1,3**	**0,1**	**0,0**	**0,4**	**0,4**	**0,7**		**1,3**	**0,3**	**-0,2**	**0,2**	**0,9**		*Seasonally adjusted domestic demand indicator 3/*

1/ Preliminar. Actualizado con información proporcionada por el Ministerio de la Producción y el INEI al 15 de julio de 2013. La información de este cuadro se ha actualizado en la Nota N° 28 (19 de julio de 2013).
2/ Incluye derechos de importación y otros impuestos.
3/ Variación porcentual respecto al mes anterior. Ajuste estacional realizado con el programa Tramo-Seats con factores estimados con información a mayo 2013.

Fuente: Ministerio de Agricultura, Ministerio de Energía y Minas, Ministerio de la Producción e Instituto Nacional de Estadística e Informática
Elaboración: Gerencia de Información y Análisis Económico - Subgerencia de Estadísticas Macroeconómicas.

nota semanal / PRODUCCIÓN / PRODUCTION

62

Notes: Although OCR accuracy is much higher than for scanned sources, gaps remain in the archive. For instance, the year 2012 is missing. Such discontinuities complicate the construction of a complete real-time dataset and require cross-checking against alternative sources.

Despite their advantages, digital sources also presented several recurring issues:

- Encoded characters replacing numerical values (e.g., `(cid:101)(cid:115)`).
- Incomplete vertical lines separating columns, creating ambiguous parsing.
- Duplicate documents (e.g., `ns-21-2015` and `ns-45-2015`), disrupting scraping routines.
- Empty table cells misinterpreted as headers (e.g., `ns-25-2014`, `ns-14-2020`).
- Missing target-period labels, complicating the assignment of growth rates (e.g., `ns-28-2013`).
- Missing year labels, complicating the chronological alignment of data (e.g., `ns-28-2013`).
- Duplicated or misaligned headers generating spurious columns.
- Double or fragmented border lines, complicating table parsing.

Figure S3 shows an example of a monthly table in digital format, where the fifth issue above is apparent: missing year labels (from May to December). While it is obvious to a human reader that the missing label corresponds to year 2012, automated scripts must explicitly account for such cases. On the other hand, Figure S4 illustrates additional issues such as incomplete borders and duplicated headers.

Figure S4. *Table 82, WR 48 (December 2016): quarterly and annual growth rates*

table 82 / cuadro 82

PRODUCTO BRUTO INTERNO /
GROSS DOMESTIC PRODUCT

(Año base 2007, Variaciones porcentuales) [1/] / *(Base year 2007, Percentage change)* [1/]

SECTORES ECONÓMICOS	2014					2015					2016			*ECONOMIC SECTORS*
	I	II	III	IV	AÑO	I	II	III	IV	AÑO	I	II	III	
Agropecuario	1,1	0,9	2,8	3,0	1,9	1,5	4,2	4,7	2,9	3,4	1,8	0,9	0,9	*Agriculture and Livestock*
Pesca	-4,8	-8,9	-15,2	-60,8	-27,9	-9,2	36,6	-21,6	43,8	15,9	1,8	-59,6	68,3	*Fishing*
Minería e hidrocarburos	4,8	-4,3	-3,0	-0,4	-0,9	4,4	7,6	10,3	15,1	9,5	15,8	23,6	15,8	*Mining and fuel*
Manufactura	3,5	-3,3	-3,7	-9,9	-3,6	-5,2	0,1	-2,0	1,2	-1,5	-2,6	-8,2	2,0	*Manufacturing*
Electricidad y agua	5,7	4,9	4,5	4,5	4,9	4,5	5,1	6,0	8,1	5,9	10,3	7,2	6,7	*Electricity and water*
Construcción	5,1	0,1	0,1	2,7	1,9	-6,8	-8,6	-6,8	-2,1	-5,8	2,1	0,9	-3,6	*Construction*
Comercio	5,2	4,4	4,0	4,2	4,4	3,6	3,8	4,1	3,9	3,9	2,8	2,3	1,4	*Commerce*
Servicios 2/	5,9	5,0	4,5	4,8	5,0	4,1	4,0	4,4	4,4	4,2	4,5	4,2	3,9	*Services*
PBI GLOBAL	**5,0**	**1,9**	**1,8**	**1,2**	**2,4**	**1,9**	**3,2**	**3,3**	**4,7**	**3,3**	**4,5**	**3,7**	**4,4**	***GDP***

1/ Preliminar. Actualizado con información al 15 de noviembre de 2016 en la Nota N° 44 (25 de noviembre de 2016).
2/ Incluye derechos de importación y otros impuestos a los productos.

Fuente: Instituto Nacional de Estadística e Informática y Ministerios de Agricultura y Riego, Energía y Minas y de la Producción.
Elaboración: Gerencia de Información y Análisis Económico - Subgerencia de Estadísticas Macroeconómicas.

nota semanal / PRODUCTO BRUTO INTERNO / GROSS DOMESTIC PRODUCT 82

Notes: While free of resolution problems, this table illustrates other challenges for automated extraction, such as irregular table borders, as the border to the right of "I" in year 2016. These features can disrupt parsing routines and require additional cleaning before integration into the RTD.

The replication codes were designed to systematically address these problems through modular, configuration-driven architecture. All cleaning operations avoid hardcoding, ensuring adaptability to format changes and preserving consistency across vintages. The modular design consists of more than 70 specialized functions organized into seven categories: text normalization (4 functions), general table cleaning (22 operations), column handling (14 functions), and table-specific transformations (13 functions each for each table).

The automated extraction workflow executes the following steps:

1. Automated downloading of the PDF WR published in the BCRP website, via Selenium-based web scraper.
2. Identification of relevant tables using keyword searches.
3. Extraction of tables with `tabula` and cleaning using specialized functions.
4. Restructuring into standardized vintage format.
5. Concatenation across vintages using union-based column alignment.
6. Export to multiple formats, such as CSV or Parquet, with comprehensive metadata.

Once both scanned and digital vintages were processed, they were integrated into a continuous RTD spanning 2000–2024. The integration workflow includes:

1. Concatenating vintages from pre- and post-2013 samples.
2. Extracting revision metadata from PDF sources.
3. Applying base-year mapping to identify structural breaks in GDP series.
4. Producing benchmark indicator datasets to flag methodological revisions.
5. Converting vintage format to release format for econometric analysis.
6. Storage of all dataset variants with comprehensive documentation.

S3 The Real-Time Dataset (RTD)

We emphasize the distinction between two key products:

- **Vintages dataset:** contains all successive records as originally extracted. This constitutes the real-time dataset (RTD).
- **Releases dataset:** derived from vintages, reorganized for revision analysis, and serving as the main input for our empirical tests.

The vintages dataset is the direct output of the extraction process and preserves the chronology of publications exactly as they appeared. By contrast, the releases dataset is a research-oriented transformation of the RTD, specifically designed to facilitate the study of revisions. Both formats are complementary: the vintages dataset ensures transparency and replicability, while the releases dataset provides the structure needed for econometric testing.

To illustrate the difference, we focus on a window that captures the lockdown effects of the COVID-19 pandemic. This period registers unusually negative monthly growth rates that are easy to track visually. Table S1 illustrates the structure of the RTD: columns represent vintages, while rows correspond to target periods or events. For instance, the July 2020 vintage contained growth figures for May 2019 through May 2020. That WR included the third estimate for March 2020 (-16.30) and the first estimate for May 2020 (-32.80). To guide interpretation, first releases in each row are highlighted in blue, second releases in red, third releases in green, and fourth releases in yellow.

The releases dataset reorganizes the same information so that each column groups all publications corresponding to the same release sequence (first, second, third, etc.) across target periods. Table S2 provides a visual representation of this reorganization. The first release of each target period is gathered in the y_t^1 column, the second in the y_t^2 column, and so forth. For example, the figures for February and April 2020 highlighted in blue in Table S1 appear in different columns within the RTD but are grouped together in the y_t^1 column of Table S2, since both correspond to first estimates. Similarly, the red-highlighted figures in Table S1 are grouped under the y_t^2 column in Table S2, and so on.

This procedure, widely adopted in the literature, makes the revision process explicit and operational. Revisions are computed as differences between consecutive release columns (e.g., $y_t^2 - y_t^1$, $y_t^3 - y_t^2$). The "Last release" column (y_t^H) contains the most recent available estimate for each target period, after which no further revisions are observed. By making the flow of updates transparent, the releases format provides a clean and interpretable structure for the rationality, efficiency, and benchmarking analyses conducted in the paper.

Table S1. *Dataset of monthly vintages*

Target period	2020m05	2020m06	2020m07	2020m08	2020m09	2020m10	2020m11	2020m12	···	2021m04	2021m05
2019m03	3.40										
2019m04	0.10	0.10									
2019m05	0.70	0.70	0.70								
2019m06	2.80	2.80	2.80	2.80							
2019m07	3.80	3.80	3.80	3.80	3.80						
2019m08	3.60	3.60	3.60	3.60	3.60	3.60					
2019m09	2.40	2.40	2.40	2.30	2.30	2.30	2.30				
2019m10	2.40	2.40	2.40	2.40	2.40	2.40	2.40	2.40			
2019m11	2.00	2.00	2.00	2.10	2.10	2.10	2.10	2.10	···		
2019m12	1.10	1.10	1.10	1.10	1.10	1.10	1.20	1.20	···		
2020m01	3.00	3.10	3.10	3.20	3.20	3.20	3.00	3.00	···		
2020m02	3.90	3.90	3.90	3.90	3.90	3.90	3.70	3.70	···	3.80	
2020m03	-16.30	-16.30	-16.30	-16.70	-16.70	-16.70	-16.30	-16.30	···	-16.80	-16.80
2020m04		-40.50	-40.50	-39.90	-39.90	-39.90	-39.20	-39.20	···	-39.10	-39.10
2020m05			-32.80	-32.70	-32.70	-32.70	-32.30	-32.30	···	-32.60	-32.60
2020m06				-18.10	-18.10	-18.10	-17.90	-17.90	···	-18.30	-18.40
2020m07					-11.70	-11.70	-11.60	-11.60	···	-11.30	-11.30
2020m08						-9.80	-9.70	-9.70	···	-9.30	-9.30
2020m09							-6.90	-6.90	···	-6.20	-6.20
2020m10								-3.80	···	-3.30	-3.30
2020m11									···	-2.50	-2.50
2020m12									···	0.50	0.60
2021m01									···	-1.00	-1.00
2021m02										-4.20	-3.80
2021m03											18.20

Notes: This table illustrates the vintage format of a real-time database. Each column corresponds to a specific data vintage — that is, the set of estimates available in the WR at the end of a given month. Each row corresponds to a target period (e.g., a specific month for which GDP growth is estimated). The publication lag of the first release is two months.

Table S2. *Dataset of monthly releases*

Target period	First release y_t^1	Second release y_t^2	Third release y_t^3	Fourth release y_t^4	···	Last release y_t^H
2020m03	-16.30	-16.30	-16.30	-16.70	···	-16.80
2020m04	-40.50	-40.50	-39.90	-39.90	···	-39.10
2020m05	-32.80	-32.70	-32.70	-32.70	···	-32.60
2020m06	-18.10	-18.10	-18.10	-17.90	···	-18.40
2020m07	-11.70	-11.70	-11.60	-11.60	···	-11.10
2020m08	-9.80	-9.70	-9.70	-9.70	···	-9.10
2020m09	-6.90	-6.90	-6.90	-6.20	···	-6.00
2020m10	-3.80	-3.80	-3.30	-3.30	···	-3.20
2020m11	-2.80	-2.50	-2.50	-2.50	···	-2.10
2020m12	0.50	0.50	0.50	0.60	···	1.00
2021m01	-1.00	-1.00	-1.00	-1.00	···	-0.80
2021m02	-4.20	-3.80	-3.80	-3.80	···	-3.60
2021m03	18.20	18.20	18.20	20.00	···	20.10

Notes: This table presents the release format of the data in Table S1, reorganized by target period to track successive estimates over time. Each row corresponds to a given target period (e.g., a specific month of GDP growth), and each column captures y_t^h, i.e. the h-th release of that estimate.

S4 Coverage of the RTD

We now summarize the coverage and structure of the RTD, which supports a comprehensive analysis of GDP as a real-time indicator of economic activity.

The dataset includes monthly, quarterly, and annual GDP growth series, with temporal and horizon coverage detailed in Table S3. Sectoral disaggregation follows INEI's level-9 classification—the most detailed level available—which distinguishes sectors such as manufacturing, commerce, agriculture, mining, and a wide range of services. The category "other services" includes real estate, education, and health.

Table S3. *Sectoral GDP as a real-time indicator*

Real-time variable (industry)	Frequency	Target period (t)	Revision horizon (h)
Real GDP	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Agriculture and livestock	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Fishing	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Mining and fuel	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Manufacturing	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Electricity and water	Monthly	2003m04–2023m10	2003m06–2024m11
	Quarterly	2003q2–2023q4	2003m08–2025m01
	Annual	2003–2023	2004m02–2025m01
Construction	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Commerce	Monthly	1993m12–2023m10	1994m02–2024m11
	Quarterly	1996q4–2023q4	1997m02–2025m01
	Annual	1996–2023	1997m02–2025m01
Other services	Monthly	1997m05–2023m10	1997m07–2024m11
	Quarterly	1997q2–2023q4	1997m09–2025m01
	Annual	1997–2023	1997m03–2025m01

Notes: The disaggregation of economic activities follows level 9 of INEI's official classification. Under this classification, the "other services" sector includes a diverse set of activities, such as real estate and personal services, including education and health (14.89%). The following categories are also identified: transportation, storage, postal and courier services (4.97%); accommodation and food services (2.86%); telecommunications and other information services (2.66%); financial and insurance services (3.22%); business services (4.24%); product taxes (8.29%); and public administration and defense (4.29%).

S5 Replication and Availability

The project is supported by a fully documented, production-ready computational pipeline designed to construct, update, and manage the RTD in a transparent and reproducible manner. All code used in this paper is publicly available under an open-source MIT license in a dedicated GitHub repository:

https://github.com/JasonCruz18/peruvian_gdp_revisions

The repository contains all `Python` codebase, configuration files, documentation, and tutorial notebooks required to reproduce the dataset construction process from raw official sources.

The guiding design principles of the pipeline are modularity, traceability, and configuration-driven execution. All data transformations are implemented as explicit, documented functions organized into specialized `Python` modules, each with a clearly defined responsibility. The main modules include: `config` (global configuration management), `scrapers` (automated data collection from official sources), `processors` (PDF processing and metadata extraction), `cleaners` (data cleaning and normalization), `transformers` (construction and transformation of real-time datasets), `orchestration` (high-level workflow coordination), and `utils` (shared utilities and record management).

The `cleaners` module plays a central role in the pipeline and contains more than 70 specialized functions designed to address the heterogeneous formatting issues present in the source tables, including text normalization, table restructuring, column alignment, numerical parsing, temporal labeling, and sector-name standardization. Crucially, no hardcoded values are used at any stage of the workflow. All pipeline behavior (scraping parameters, table parsing rules, cleaning tolerances, metadata handling, and output formats) is controlled through human-readable configuration files in `config/config.yaml`.

Replication relies on the execution of the single `python scripts/update_rtd.py`, generating the complete workflow from raw inputs to final vintage and release datasets. Internally, this script orchestrates specialized runners for different data sources and table types, ensuring a unified and reproducible execution path across historical and recent vintages. The pipeline is idempotent: previously processed vintages are automatically detected and skipped through record-tracking utilities implemented in `utils`. This design allows users to update the RTD incrementally as new publications become available, without reprocessing the full historical archive. Execution logs and metadata files are generated at each stage, enabling full auditability of the construction process.

The pipeline relies exclusively on official public data sources. GDP growth figures are produced by the INEI and disseminated through the publications of the BCRP. For the post-2013 period, inputs are obtained automatically via web-scraping routines implemented in `scrapers`. For earlier periods, replication relies on cleaned intermediate datasets derived from historical publications, together with the full set of extraction, cleaning, and validation routines documented in the repository. As discussed in previous sections, legal and practical restrictions prevent redistribution of some original scanned documents, but these constraints do not affect the reproducibility of the results.

To accommodate different analytical needs, the pipeline produces multiple dataset variants. The core output is the RTD in vintage format, constructed to preserve the exact chronology of official publications as they appeared in real time. From this representation, the pipeline generates the release-format datasets used in the empirical analysis, where estimates are reorganized by release order to facilitate the computation of revisions and the implementation of rationality and efficiency tests. In addition, the pipeline produces auxiliary datasets that explicitly account for methodological changes in the official statistics. These include benchmark-indicator datasets that flag coordinated revisions associated with rebasing or definitional updates, as well as base-year-adjusted datasets that mark observations affected by structural breaks. Rather than removing affected observations, the pipeline applies explicit sentinel values and metadata flags, preserving the full revision history while allowing researchers to condition their analysis on comparability constraints.

All datasets are exported in both CSV and Parquet formats. The CSV files provide maximum accessibility and ease of inspection, while the Parquet files preserve data types and offer improved performance for large-scale analysis. Each dataset is accompanied by machine-readable metadata describing variable definitions, release horizons, base-year adjustments, and benchmark indicators, ensuring interoperability with a wide range of statistical software environments.

To facilitate replication and reuse by non-specialist users, the repository includes a set of fully documented Jupyter notebooks located in the `notebooks/` directory that illustrate each stage of the workflow. These notebooks provide executable examples covering data ingestion, cleaning, RTD construction, and release-format conversion, together with explanatory text and intermediate outputs. The project also includes an optional interactive dashboard implemented using Streamlit (`dashboard/app.py`), which allows users to explore the RTD visually, navigate across vintages and release horizons, inspect revision patterns, and export filtered datasets without writing code. While not required for replication of the paper's results, the dashboard facilitates exploratory analysis and lowers the barrier to entry for applied users.

Reproducibility is further ensured through explicit dependency management and version control. The repository provides pinned dependency files that allow users to recreate the exact computational environment used in the paper. All code is versioned using semantic versioning, and stable releases are tagged on GitHub. Automated tests are executed through continuous integration workflows across operating systems and `Python` versions, ensuring cross-platform stability of the pipeline.

To promote long-term accessibility, a stable release of the codebase and associated datasets will be archived and assigned a permanent DOI upon acceptance of the paper. Researchers using the RTD or the replication code are encouraged to cite both the published article and the archived software release, following the citation instructions provided in the repository. The combination of open-source code, documented workflows, and explicit data availability statements ensures that the empirical results of this paper meet high standards of transparency, reproducibility, and extensibility.