The Seasons They Are A-Changin': A Century of Definitions and a Way Forward*

Authors: Carter Bryson and Gary Cornwall, U.S. Bureau of Economic Analysis†
Date: March 26, 2026
Abstract: What is seasonality? To date, there is not yet a consensus definition, as over a dozen unique definitions of seasonality can be found in the economics and statistics literature. This lack of agreement complicates the identification of, and adjustment for, seasonal effects in economic time series. In this paper, we propose a new definition of seasonality as well as how to detect and adjust for it, ultimately aiming to build consensus around a more unified approach. We first review the literature and identify common themes with an eye towards formulating a foundational definition – expressed in both plain language and mathematical terms – that can be used by academics, data providers, policymakers, and practitioners alike. We then propose to classify a series as seasonal if, among all its possible repeating cycles, the ones associated with the fundamental periodicity are demonstrably larger than the others. In other words, we define seasonality as a measure of relative peak dominance in the spectral density function, following Nerlove (1964) and Granger (1978). Next, we develop a new methodology for detecting seasonality and performing seasonal adjustment that follows directly from our definition and leverages properties of the sample periodogram, the empirical counterpart to the spectral density function. We illustrate the performance of our test statistic through simulation studies and find that it is well-sized and has good power. Lastly, we introduce Stochastic Spectral Imputation (SSI), an adjustment procedure that outperforms standard methods currently used by national statistical offices when applied to prominent macroeconomic time series.
Keywords: Time-Series, Seasonality, Seasonal Adjustment, Frequency Domain
JEL Codes: C22, C12, C15, E37



*The views expressed in this paper are those of the authors and not necessarily those of the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce.

†Contact information: carter.bryson@bea.gov and gary.cornwall@bea.gov. Address: U.S. Department of Commerce, Bureau of Economic Analysis, 4600 Silver Hill Road, Suitland, MD 20746

1. Introduction

Seasonality is among the most recognizable features of economic time series. In many settings it is visible to the naked eye, yet the field has never converged on a single, operational definition of what it means for a series to be "seasonal." This plurality has legitimate origins. The phenomena economists call "seasonal" are themselves heterogeneous: deterministic calendar effects, stochastic seasonal dynamics, seasonal unit roots, and slowly evolving periodic patterns are genuinely different objects and internally consistent formalizations exist for each.[1] The diagnostic suites and recommended workflows that accompany production pipelines (most prominently X-13ARIMA-SEATS and TRAMO-SEATS) were designed to navigate this heterogeneity.[2] The resulting patchwork of definitions, tests, and adjustment procedures reflects both the genuine heterogeneity of seasonal phenomena and the organic co-evolution of software and methods to capture them. The accumulation of many such tools and concepts is a rational response to a complex empirical reality.[3]

Consider, however, the structure of this accumulated toolkit from the perspective of set theory. Among all time series, there is a subset whose members exhibit seasonality. The existing battery of tests (seasonal dummies, HEGY, Canova-Hansen, spectral diagnostics, autocorrelation-based procedures) each targets a *strict subset* of this seasonal partition: deterministic seasonal means, seasonal unit roots, seasonal stability, spectral peaks, and so on.[4] These subsets partially overlap but are not equivalent; a series may be seasonal under one test and nonseasonal under another, not because either test is deficient, but because they are designed for different population objects. The field's current practice therefore answers "what *type* of seasonality?" before resolving the logically prior question: "is the series in the seasonal subset at all?"

Against this backdrop, we argue that addressing the superset question, particularly in an end-to-end fashion as is needed by national statistical offices, must satisfy three properties simultaneously: (1) an explicit population target, (2) a test statistic with known and tractable distributional properties for that target, and (3) an adjustment operator that is the logical conjugate of the definition and test. Individual elements of this triad can be found in prior work: Nerlove (1964) and Granger (1978) proposed spectral population objects; Hylleberg et al. (1990) and Canova and Hansen (1995) developed tests with well-characterized null distributions; Maravall (2011) and the X-13 pipeline implement adjustment operators that are optimal under their respective model assumptions. Our contribution is the *simultaneous* satisfaction of all three criteria within a single closed-loop chain. Specifically, we define seasonality as a form of *relative peak dominance*

[1] See Falkner (1924); Kuznets (1933); Nerlove (1964); Granger (1978); Hylleberg et al. (1990); Canova and Hansen (1995); Ghysels and Osborn (2001).

[2] These diagnostic tools and recommended workflows are further embedded in international guidance intended to promote quality, coherence, and transparency. For example, the ESS Guidelines on Seasonal Adjustment emphasize coherent practice, metadata, and transparency across domains and producers (Commission and Eurostat, 2015, 2024). The U.S. Census Bureau's X-13ARIMA-SEATS documentation similarly emphasizes diagnostics for residual seasonality and stability. See Dagum (1980); Findley et al. (1998); Gómez and Maravall (2001); Maravall (2011); Lytras et al. (2007); McElroy and Roy (2022) as well.

[3] The link between individual diagnostics and an explicit object of interest representing what is "seasonal" is often indirect; see Dagum (1980); Lytras et al. (2007); Maravall (2011); Webel and Ollech (2018); Chen and Cornwall (2021) for discussion.

[4] See Dagum (1980); Findley et al. (1998); Gómez and Maravall (2001); Maravall (2011); Lytras et al. (2007); McElroy and Roy (2022).

in the spectral density function: a series is seasonal if the height of the spectral density within a prespecified set of "seasonal" frequency intervals exceeds the height in the remaining intervals. This yields a population object that formalizes seasonality as a comparative statement about two *ex ante* frequency subsets, following Nerlove (1964), who first identified spectral peaks at seasonal frequencies as the defining characteristic of seasonality, and Granger (1978), who formalized the population-sample distinction for spectral seasonality.[5]

Working in the frequency domain is not merely a technical convenience; it is motivated by four substantive considerations. First, the frequency domain characterizes periodic behavior by repetition rate rather than calendar position, making the resulting definition invariant to the choice of calendar (Gregorian, fiscal, lunar, or otherwise). Second, a frequency-domain formulation does not presuppose the *type* of seasonal process generating the data; it targets a superset property (excess spectral mass at seasonal frequencies) that is shared by deterministic, stochastic, and unit-root seasonal processes alike. Third, the spectral density function provides a natural variance decomposition by frequency, so that seasonality becomes a statement about the *disproportionate* contribution of seasonal frequencies to total variance, rather than a statement about mean structure or autocovariance at specific lags. Fourth, effects that complicate time-domain definitions (trading-day variation, outliers, and moving holidays) are treated as nuisance effects in the frequency domain. Trading-day effects and outliers affect spectral estimation but not the definition itself; moving holidays, which migrate across the calendar, do not concentrate power at fixed seasonal frequencies and are therefore correctly scoped out as calendar effects rather than periodic structure. In practice, such effects are addressed through standard pre-filtering prior to spectral analysis.

From this definition and population object, we derive a finite-sample test statistic based on the sample periodogram. The periodogram is a well-known inconsistent estimator of the spectral density function, and we exploit this inconsistency directly: periodogram ordinates at Fourier frequencies are approximately independently exponentially distributed (Priestley, 1981), a property that would not hold under consistent estimation. We apply results from extreme value theory to show that our test statistic converges to a Logistic distribution with known location and scale parameters. The test therefore requires no simulation-based critical values, bandwidth selection, or kernel specification. We then propose an adjustment method, *Stochastic Spectral Imputation* (SSI), that is definitionally consistent with our target object and operates on the discrete Fourier transform of the observed series. Rather than mechanically suppressing frequencies in the seasonal intervals as in notch filtering, SSI replaces their magnitudes with draws from an empirically estimated "donor" distribution constructed from the nonseasonal intervals. Phase can be optionally kept or randomized and conjugate symmetry is enforced accordingly. The resulting adjusted series, by construction, removes exactly the excess peak power that defines seasonality in our framework while preserving the nonseasonal spectral floor. The identification of the mismatch between a population object, the Wiener-Kolmogorov filter that estimates it, and the implicit seasonal component in production pipelines has a long lineage, beginning with Nerlove (1964); SSI operationalizes one resolution of that mismatch. Because SSI is stochastic (each application

[5]The framework applies to weakly stationary processes. Extension to the non-stationary case proceeds through standard pre-filtering (e.g., differencing); the contribution here is the closed-loop structure linking definition, test, and adjustment, not the pre-filtering step.

yields a different adjusted series), production implementation may average over replications or report uncertainty bands; we discuss these options in Section 4.

National statistical offices and major data providers are necessarily resource constrained: they must identify and seasonally adjust large, evolving inventories of series; maintain production schedules; support diverse user communities; and document methods in a way that is both auditable and communicable.[6] In such an environment, properties such as statistical robustness, analytic transparency, minimal parametric commitment, and scalability are not luxuries; they are design requirements (Commission and Eurostat, 2024; United Nations et al., 2025). Our proposed methodology, which includes a definition, test statistic, and adjustment procedure, fulfills these requirements, making it an ideal framework around which data providers and national statistical offices can coalesce.

We view the definition-test-adjustment chain developed in this paper as a proof of concept that epistemic closure in seasonal adjustment is achievable. We do not claim that our frequency-domain definition is the only valid formalization of seasonality; the type-specific tests and procedures developed over the past century remain valuable for characterizing *which kind* of seasonality a series exhibits. Rather, our claim is narrower and, we believe, stronger: once a definition of seasonality is adopted, the downstream test and adjustment should be its logical consequences, and the three elements should form a closed system in which each step is derivable from the preceding one. We demonstrate that this is feasible in the frequency domain. We challenge others to construct equivalent closed-loop chains in their preferred domains (time-domain, state-space, or otherwise), thereby establishing a common architecture and platform for communication even if the specific definitions differ.

The remainder of the paper proceeds as follows. Section 2 reviews prominent definitions of seasonality and how they map (often imperfectly) into diagnostic and adjustment practice. Section 3 formalizes our population object and derives the sample statistic and its limiting null distribution. Section 4 presents SSI and establishes its definitional consistency and asymptotic justification. Section 5 studies finite-sample size, power, and adjustment properties under a range of seasonal and nonseasonal data-generating processes. Section 6 illustrates the testing procedure and adjustment operator on established time series with comparison of the final product to published seasonally adjusted series. Section 7 discusses implications for official statistics production, documentation, and future extensions.

2. A Century of Discourse on Seasonality

Seasonality has been at the center of econometrics and the analysis of economic time series going back to at least Falkner (1924), who argued that observed movements can be decomposed

[6] For example, the Bureau of Economic Analysis (BEA) is primarily known for producing a measure of Gross Domestic Product and its components which make up the National Income and Product Accounts (NIPA). However, the BEA also provides supplemental measures of the economy in satellite accounts (e.g., space economy) and all told produces approximately four million time series measuring economic activity. Therefore, a statistically robust, analytically transparent, scalable, and epistemologically consistent framework would not only benefit the agency and its stakeholders, but also serve to be an efficient use of taxpayer dollars in pursuit of their stated mission from Congress.

into a secular trend, cyclical fluctuations, irregular deviations, and seasonal variation. While the intervening decades have produced an expansive methodological literature, it remains conventional, especially in official statistics, to describe an observed series as a composition of trend-cycle, seasonal, and irregular components.[7] This section presents a taxonomy of how seasonality has been defined and operationalized over time, organized around the population objects, test hypotheses, and adjustment operators that different approaches employ. The goal is not to adjudicate which historical approach is "right"; for specific seasonal types (fixed effects, stochastic cycles, unit roots), rigorous and well-characterized objects, tests, and procedures exist individually. Rather, the purpose is to clarify where these approaches overlap and where they diverge, and to identify the superset question that the accumulated literature leaves open: what does it mean for a series to be "seasonal" at all, prior to any commitment about type? The field has produced multiple coherent but non-equivalent notions of "seasonality," and these notions are routinely combined within the same workflow without an explicit closure linking definition, test, and adjustment.

2.1. Definitions

First, a common theme across definitions is that of regularity and/or repetitiveness. Falkner (1924) defined seasonal variation as "...part of the fluctuation due to the persistent tendency for certain months of each year to be regularly higher than certain other months of the year...". Many decades later, the U.S. Census Bureau described seasonality in similar language: "regularly periodic fluctuations which recur every year with about the same timing and same intensity" (Kallek, 1978). The broad intuition is stable across time: seasonal behavior is an intra-cycle pattern that repeats. Yet the technical commitments differ. Falkner's language is restrictive (monthly measurement; directional emphasis), whereas later official-statistics language is more permissive (unspecified frequency; ambiguous direction; implicit accommodation of trading-day and holiday effects through calendar adjustment).[8]

Second, authors often distinguish between deterministic and stochastic seasonality without addressing the more fundamental question of what constitutes seasonality in the first place. For example, a modern, intuitive characterization is that deterministic seasonality requires a stable pattern, such as a sinusoid with time-invariant amplitude and frequency, whereas stochastic seasonality permits the pattern to evolve over time, such as a sinusoid with time varying amplitude and/or frequency (Hylleberg et al., 1990; Canova, 1992; Franses, 1994; Canova and Hansen, 1995; Busetti and Harvey, 2003; Ghysels et al., 2006; Doppelt, 2023). In either case, the presence of seasonality is predicated on agreement that inclusion of a sinusoid, regardless of its properties, makes the series seasonal in nature. Moreover, this dichotomy is deeply ingrained in contemporary testing practice. In many cases, test statistics are designed not to detect the presence of some underlying notion of seasonality, but rather to distinguish between different *types* of seasonality, deterministic

[7] Today this decomposition is often introduced in the context of seasonal *adjustment* rather than seasonal *definition*. For example, Ghysels and Osborn (2001) introduces deterministic and stochastic seasonal processes before presenting standard decomposition structures (Chapter 4). In official-statistics practice, the same decomposition language is used across filter-based and model-based pipelines, even though the induced notions of "seasonal" differ across methods.

[8] The apparent similarity of prose across definitions can obscure deep differences in the underlying primitives. In particular, some definitions are fundamentally about the conditional mean (calendar effects), others about persistence (dependence across annual lags), and others about localized spectral structure. These objects coincide only under special conditions.

versus stochastic (Hylleberg et al., 1990). This procedure presupposes that the series is seasonal to begin with.

Definitions in the literature also differ with respect to the *domain* in which they are formalized. Time series admit two complementary representations: the time domain, tied to sequential ordering and a calendar, and the frequency domain. Note that "calendar" is not unique: Julian, Gregorian, lunar, and fiscal calendars induce different partitions of time into seasons. A definition that is invariant to calendar choice is therefore conceptually distinct from one that is not. Conversely, the frequency domain is tied to repetition rates rather than the natural ordering of events. Some definitions translate cleanly across domains; others do not.

For example, definitions cast in the time domain typically operationalize seasonality by referencing a year (or a within-year partition) as the primitive marker. Ghysels et al. (2006) defines seasonality as a periodic and largely repetitive pattern observed *over the course of a year*. Other authors emphasize "variation within the year" (Sims, 1993) or evolution of within-year patterns (Jaditz, 1994). A more pragmatic line defines seasonality through time domain indicators such as the autocorrelation function: a series is seasonal if correlations at seasonal lags exhibit systematic peaks or troughs (McElroy, 2008).[9] These approaches implicitly take the calendar partition (often annual) as given and define seasonality via conditional mean structure, stability of seasonal coefficients, or dependence at seasonal lags.

Yet restricting attention to annual partitions is not innocuous. Economic behavior is not confined to the Gregorian annual cycle: religious calendars (e.g., Ramadan in the Hijri calendar), election cycles, and institutional schedules generate periodicities that may migrate across the Gregorian year or occur at other horizons. Whether such phenomena "count" as seasonality depends on the primitive: a calendar-anchored definition will treat migrating lunar holidays differently than a definition based on repetition rates.

On the other hand, the frequency domain characterizes periodic behavior by how often it repeats rather than by where it falls on a particular calendar. Nerlove (1964) was the first to define seasonality explicitly as the characteristic that gives rise to spectral peaks at seasonal frequencies, thereby establishing both the chronological and intellectual priority of spectral definitions. Building on Nerlove's insight, Granger (1978) formalized an exact population and sample distinction, sharpening the argument that seasonality is most naturally defined in spectral terms:

> **Definition 1:** The process X_t has property S if $f(\omega)$ has peaks within $\omega_s(\delta)$ for some small $\delta > 0$.
> **Definition 2:** The series x_t apparently has property S if $\hat{f}(\omega)$ has peaks in $\omega_s(\delta)$ for some small $\delta > 0$.

This formulation is notable because it explicitly separates the population object $f(\omega)$ from the sample diagnostic $\hat{f}(\omega)$. However, the primitive notions remain under-specified: what constitutes a peak (local vs global, relative to what baseline), and how small or large should δ be? Subsequent work sought to clarify these primitives. Soukup and Findley (1999) proposed a "visual significance" notion of peaks based on differences of neighboring ordinates of a transformed spectrum, calibrated

[9] Allowing troughs corresponds to permitting negative seasonal dependence, which can arise naturally in seasonal AR structures and in interactions across seasonal lags.

by empirical thresholds. More recent work by McElroy and Roy (2021); McElroy (2021) and the comprehensive review in McElroy and Roy (2022) represents substantial progress along this line, developing formal statistics for inference and adequacy assessment with rigorous distributional theory.[10] Two challenges have constrained the broader adoption of frequency-domain definitions and testing: (i) the difficulty in operationalizing the primitive notions in a formal manner (e.g., δ), and (ii) the difficulty in obtaining an estimate of the spectral density function, $\hat{f}(\omega)$, with desired statistical properties such as consistency and efficiency.

2.2. Tests and Adjustment

Because definitions imply distinct population and sample objects, tests for seasonality necessarily encode different null–alternative pairs. Table 1 re-emphasizes that common procedures do not test a single generic hypothesis "seasonality exists," but rather test object-specific statements: seasonal means versus equality, stability versus drift, stationarity versus seasonal integration, or mild versus strong persistence in designated frequency neighborhoods. The nulls/alternatives are schematic and the key point to take away is that each procedure targets a different object (means, stability, roots, localized spectral behavior), so "seasonal" conclusions are not automatically comparable across tests. Put differently, the plurality of tests is a direct consequence of the plurality of population objects.

[Insert Table 1 Here]

This observation clarifies the inferential structure of seasonal detection in practice. Because individual tests target different population objects, applied workflows frequently become diagnostic assemblages: researchers or agencies consult multiple procedures until the totality of evidence supports a decision to adjust or not adjust. This practice is rational given the heterogeneity of seasonal phenomena, and it has real benefits (flexibility to handle diverse seasonal types, robustness to specification error in any single test, and adaptability to the wide range of series encountered in production). Nonetheless, the assemblage strategy also carries a cost: without an explicit statement of which superset object is being tested, the inferential content of "the series is seasonal" remains conditional on the particular combination of diagnostics employed.

Seasonal adjustment necessarily operationalizes a definition of seasonality because it implements a transformation that removes something. Filter-based pipelines in the X-11 tradition define "seasonal" as what is extracted by sequences of moving-average filters (Kallek, 1978), with options determining the effective pass-bands and the degree of symmetry in the implied frequency response. Model-based pipelines (SEATS/TRAMO-SEATS) define "seasonal" through ARIMA-based signal extraction, where the seasonal component and the adjustment filter are determined by the fitted decomposition structure. In both cases, the induced notion of seasonality is a property of the adjustment mechanism: seasonality is what the filter (or the model-implied Wiener–Kolmogorov filter) removes.

The taxonomy clarifies why the prevailing assemblages of definitions, tests, and adjustments can

[10] This line of work is framed in the context of post-adjustment diagnostics, that is the ability to detect if an adjusted series exhibits residual seasonality. This framing is not accidental: in production environments, the relevant question is frequently whether an adjustment pipeline has removed sufficiently much of whatever is meant by "seasonal" under the adopted diagnostics. The McElroy and Roy (2022) review, in particular, provides a systematic mapping between diagnostic procedures and the spectral properties they target.

exhibit a systematic gap, even when each element works well individually and in combination across a wide range of production settings. A mean-contrast test does not uniquely rationalize a particular seasonal filter. A seasonal unit root test does not uniquely rationalize a symmetric smoothing pipeline. A spectral-peak diagnostic does not uniquely rationalize a moving-average decomposition unless the adjustment is explicitly designed as an intervention on the same spectral object that generated the diagnostic signal. Consequently, even when a test is formally correct relative to its own object, the subsequent adjustment step may target a different object. The practical consequences of this gap are most visible not within any single agency's pipeline (where accumulated institutional knowledge often bridges the disconnect effectively) but in cross-agency communication and comparability: when two producers adjust the same series under different definition-test-adjustment assemblages, the resulting "seasonally adjusted" outputs may not be directly comparable, and the source of disagreement is structural rather than parametric.

The historical record therefore supports two simultaneous claims. First, the literature contains multiple coherent and practically useful notions of seasonality, along with sophisticated tests and highly effective production pipelines. Second, as a field-level practice, seasonality remains epistemically inconsistent because these elements are not generally assembled into a single closed system. Definitions are plural; tests inherit that pluralism via distinct null–alternative pairs; adjustment pipelines define seasonality operationally through their own mechanisms; and the link between the inferential target and the transformation target is rarely made explicit.

The consequences of this structure are practical as well as conceptual, and they cut in both directions. On one hand, the plurality of tools gives agencies genuine flexibility: they can tailor pipelines to the specific characteristics of individual series, accommodate diverse seasonal types within a single inventory, and draw on accumulated institutional expertise that bridges gaps between formal diagnostics and production decisions. On the other hand, within and between national statistical offices, "seasonally adjusted" can describe outputs generated under materially different targets and diagnostics, complicating comparability and revision narratives. For users, the absence of an explicit superset population object makes it difficult to interpret disagreements across providers, to reproduce adjustments, or to evaluate adequacy beyond a collection of heterogeneous flags. For production-scale implementation, the reliance on multi-diagnostic assemblages and option-dependent pipelines can limit transparency and make scalability contingent on conventions and expert overrides that are difficult to communicate to external stakeholders.

These observations motivate the closed-loop alternative developed in subsequent sections. The central design principle is epistemic closure: define a frequency-domain population object via an explicit partition of seasonal and nonseasonal subsets; construct a test statistic whose null is tied to that same object; and implement adjustment as a targeted intervention that modifies exactly the subset implicated by the test. Our claim is not that prior definitions or pipelines were misguided; as the preceding discussion makes clear, the accumulated toolkit reflects a rational response to genuinely heterogeneous phenomena. Rather, the claim is that this accumulated pluralism has left a specific gap: the superset question of whether a series is seasonal at all, prior to any commitment about type. The definition-test-adjustment chain we develop is a proof of concept that this gap can be closed. Type-specific tests remain valuable for characterizing *which kind* of seasonality a series exhibits; what our framework adds is a logically prior layer that addresses the membership question for the seasonal set as a whole.

3. A Way Forward: Definition and Test Statistic

In this section we pursue two objectives that mirror the themes of Section 2. First, we define seasonality in the frequency domain through a population object that makes its primitive notions explicit. Second, we derive a sample statistic and a null distribution that are tied to that same object, thereby enforcing the epistemic closure that is often absent in prevailing definition–test–adjustment workflows.

Our approach rests on two building blocks. The first is a fixed *frame of reference*: a deterministic partition of the frequency domain into disjoint bins, chosen *ex ante*. The second is an operational definition of a *peak* as a relative notion of height, defined by extrema computed within the same bins. This construction abstracts away from calendar-specific regressors (Easter, trading-day, moving holidays) and instead targets repetitive behavior as a property of spectral structure, while retaining transparency about the subjective inputs that remain—namely the designation of a seasonal frequency set. The contribution is not to eliminate all discretion, but to narrow the scope for idiosyncratic choices by anchoring both testing and (later) adjustment to the same explicitly defined subset of the spectrum.

3.1. A Population Object Defined by a Seasonal Frequency Partition

Let $\{Z_t\}_{t\in\mathbb{Z}}$ be a real-valued, weakly stationary process with mean μ, autocovariance function

$$\gamma_Z(\tau) \;=\; \mathbb{E}[(Z_t-\mu)(Z_{t+\tau}-\mu)]\,, \tag{1}$$

and spectral density

$$f_Z(\omega) \;=\; \sum_{\tau=-\infty}^{\infty}\gamma_Z(\tau)e^{-i\omega\tau}, \qquad \omega\in[-\pi,\pi], \tag{2}$$

with the inversion formula $\gamma_Z(\tau)=(2\pi)^{-1}\int_{-\pi}^{\pi} f_Z(\omega)e^{i\omega\tau}\,d\omega$. We impose the standard regularity condition $\sum_\tau|\gamma_Z(\tau)|<\infty$, implying that f_Z is continuous on $[-\pi,\pi]$. We emphasize that the population object in this section is f_Z, not a decomposition component, seasonal dummy coefficients, or a seasonal unit root factor; the taxonomy in Section 2 clarifies that these are distinct objects, and our goal here is to commit to one explicitly.

Defining Primitives Let $P_{\max}\in\mathbb{N}$ be an *ex ante* upper bound on the longest fundamental period (in years) we are willing to regard as seasonally relevant, and let N be the number of observations per calendar year. For each fundamental p define the associated harmonic set

$$\Omega_p(N)=\left\{\frac{2\pi k}{pN}\;:\;k=1,2,\ldots,\left\lfloor\frac{pN}{2}\right\rfloor\right\}\subset(0,\pi), \tag{3}$$

which excludes $\omega=0$ and (when pN is even) excludes the Nyquist frequency π. The *complete* candidate set of calendar harmonics up to $P_{\max}$ years is

$$\Omega(N;P_{\max})=\bigcup_{p=1}^{P_{\max}}\Omega_p(N). \tag{4}$$

Given a chosen set of fundamentals $\mathcal{P} \subseteq \{1, \ldots, P_{\max}\}$ (e.g., $\mathcal{P} = \{1, 2, 4\}$ for annual, biennial, and quadrennial cycles), the induced "Gregorian" seasonal subset is

$$\Omega_G(\mathcal{P}, N) = \bigcup_{p \in \mathcal{P}} \Omega_p(N) \subseteq \Omega(N; P_{\max}). \tag{5}$$

The common special case is $\mathcal{P} = \{1\}$, but longer fundamentals ($p > 1$) can be empirically relevant (e.g., quadrennial political cycles) and to avoid any degeneracy it is assumed $\mathcal{P} \neq \emptyset$.

Partitioning the Frequency Domain Rather than introducing *ad hoc* windows $\omega_s \pm \delta$ around points in $\Omega_G(\mathcal{P}, N)$ (Granger, 1978), we fix a deterministic partition of $(0, \pi)$ into M disjoint, equal-width bins,

$$B_m = \left(\frac{\pi(m-1)}{M}, \frac{\pi m}{M} \right], \qquad m = 1, \ldots, M, \tag{6}$$

with bin maximum,

$$b_m(f_Z) := \sup_{\omega \in B_m} f_Z(\omega), \tag{7}$$

and classify bins according to whether they intersect the seasonal set:

$$\begin{aligned} \mathcal{M}_1(\mathcal{P}, N; M) &= \{m : \mathrm{int}(B_m) \cap \Omega_G(\mathcal{P}, N) \neq \emptyset\} \\ \mathcal{M}_0(\mathcal{P}, N; M) &= \{1, \ldots, M\} \setminus \mathcal{M}_1(\mathcal{P}, N; M). \end{aligned} \tag{8}$$

We treat M as a design parameter fixed *ex ante*. Since M is fixed, $T/M \to \infty$ as $T \to \infty$, ensuring that the number of Fourier frequencies per bin diverges. We also choose M so that each $\omega_s \in \Omega_G(\mathcal{P}, N)$ lies in the interior of some bin B_m; this is always possible because, for any fixed ω_s, the set of M placing ω_s on a bin boundary is countable, hence avoidable. The conceptual role of (6)–(8) is to render the frame of reference fixed *ex ante*: the bins are declared before inspecting the data, and the only remaining modeling input is $\Omega_G(\mathcal{P}, N)$, which is itself explicit and interpretable.

The Population Object We define a population separation functional

$$\Delta(f_Z; \mathcal{P}, N, M) = \underbrace{\max_{m \in \mathcal{M}_1(\mathcal{P}, N; M)} b_m(f_Z)}_{\text{largest seasonal-bin peak}} - \underbrace{\max_{m \in \mathcal{M}_0(\mathcal{P}, N; M)} b_m(f_Z)}_{\text{largest nonseasonal-bin peak}}. \tag{9}$$

We call $\{Z_t\}$ *seasonal relative to* $(\mathcal{P}, N, M)$ if $\Delta(f_Z; \mathcal{P}, N, M) > 0$. This definition is intentionally narrow: it declares seasonality as a *relative* property of spectral maxima across two explicitly defined subsets of the frequency domain. It thereby differs from mean-based definitions and from unit-root definitions in the taxonomy, and it is designed to align with the operational logic of spectral-peak diagnostics while making the underlying primitives (the partition and the peak operator) explicit.

Although motivated by "peaks" language, $\Delta(f_Z; \mathcal{P}, N, M)$ is an extreme-value contrast of upper envelopes across two frequency subsets. This includes both narrow-band seasonal peaks and diffuse seasonal energy that remains concentrated within the seasonal bins. The definition is

intentionally indexed by the design resolution M; it formalizes a particular notion of seasonality, dominance of seasonal-bin spectral power relative to the nonseasonal envelope, chosen to align exactly with the sample test statistic and adjustment procedure introduced below.

Two remarks clarify the role of M. First, M governs the resolution of the frame of reference: the bin width π/M determines how localized the notion of a "seasonal peak" is around the calendar harmonics in $\Omega_G(\mathcal{P}, N)$. Smaller bins (larger M) yield a more localized comparison, while larger bins (smaller M) aggregate across wider frequency ranges; in either case, the definition remains anchored to the same *ex ante* partition through $b_m(f_Z)$ and $\mathcal{M}_1(\mathcal{P}, N; M)$. Second, the asymptotic requirement $T/M \to \infty$ is imposed only to support the sample analog developed below: it ensures that each bin contains a diverging number of Fourier frequencies so that within-bin maxima of the periodogram admit classical extreme-value approximations. Accordingly, $\Delta(f_Z; \mathcal{P}, N, M)$ should be interpreted as a population object indexed by the fixed design choice M, rather than as a limit in M that collapses the seasonal set to points.

3.2. A Sample Statistic and Its Limiting Distribution

We now construct a statistic that targets the same object as that in Equation 9. Let $\{z_t\}_{t=1}^T$ be a finite sample from $\{Z_t\}$, and let $\bar{z} = T^{-1}\sum_{t=1}^T z_t$. Define the discrete Fourier frequencies $\omega_j = 2\pi j/T$ for $j \in \mathcal{J} := \{1, \ldots, \lfloor (T-1)/2 \rfloor\}$. Without loss of generality, we work with the demeaned sample, $\tilde{z}_t = z_t - \bar{z}$ and define the periodogram using $\tilde{z}_t$. For $j \geq 1$, this replacement does not affect the asymptotic distribution under the null. The (raw) periodogram is

$$I_T(\omega_j) = \frac{1}{T}\left|\sum_{t=1}^T \tilde{z}_t e^{-i\omega_j t}\right|^2, \qquad j \in \mathcal{J}. \tag{10}$$

As is well known, the sample periodogram is not a consistent estimator of $f_Z(\omega)$ for fixed ω. However, rather than attempting to correct for this inconsistency, we exploit it: under a white-noise null, the periodogram ordinates have a stable, non-degenerate limiting distribution and, crucially, their maxima are characterized by classical extreme-value approximations.

Deriving the Null Distribution Under H_0 we assume Z_t is white noise with variance σ^2, so $f_Z(\omega) \equiv \sigma^2$ on $(0, \pi)$ and, for $j \in \mathcal{J}$, the periodogram ordinates satisfy $I_T(\omega_j)/\sigma^2 \Rightarrow \mathrm{Exp}(1)$ and are asymptotically independent (Politis and McElroy, 2019), where $\mathrm{Exp}(1)$ denotes an exponential with unit rate (equivalently, unit mean). The alternative of interest is the class of processes satisfying $\Delta(f_Z; \mathcal{P}, N, M) > 0$.

Using B_m as defined earlier, let $\mathcal{J}_m := \{j \in \mathcal{J} : \omega_j \in B_m\}$ and $n_m := |\mathcal{J}_m|$. Define the within-bin periodogram maximum as

$$b_m(I_T) := \max_{j \in \mathcal{J}_m} I_T(\omega_j). \tag{11}$$

Then, we can define the sample statistic as

$$\widehat{\Delta}(z; \mathcal{P}, N, M) := \max_{m \in \mathcal{M}_1(\mathcal{P}, N; M)} b_m(I_T) - \max_{m \in \mathcal{M}_0(\mathcal{P}, N; M)} b_m(I_T). \tag{12}$$

Because the bins form a partition, we have that $\max_{m \in \mathcal{M}_1} b_m(I_T) = \max_{j \in \mathcal{J}_1} I_T(\omega_j)$ and similarly for $\mathcal{M}_0$, where

$$\mathcal{J}_1 := \{j : \omega_j \in B_m \text{ for some } m \in \mathcal{M}_1(\mathcal{P}, N; M)\}, \qquad \mathcal{J}_0 := \mathcal{J} \setminus \mathcal{J}_1. \tag{13}$$

Let $N_1 := |\mathcal{J}_1|$ and $N_0 := |\mathcal{J}_0|$ denote the total count of ordinates in seasonal bins and non-seasonal bins accordingly. Under H_0, $I_T(\omega_j)/\tau \Rightarrow \mathrm{Exp}(1)$ with asymptotic independence across distinct Fourier frequencies, where $\tau = \mathbb{E}[I_T(\omega_j)] = \sigma^2$. Classical extreme-value theory (EVT) for i.i.d. exponential distributions implies that, for $g \in \{0, 1\}$,

$$\frac{\max_{j\in\mathcal{J}_g} I_T(\omega_j) - \tau \log N_g}{\tau} \;\Rightarrow\; \mathrm{Gumbel}(0, 1), \tag{14}$$

so $\max_{j\in\mathcal{J}_g} I_T(\omega_j) \approx \mathrm{Gumbel}(\tau \log N_g, \tau)$. Moreover, the maxima over $\mathcal{J}_1$ and $\mathcal{J}_0$ are (asymptotically) independent since they are functions of disjoint sets of ordinates. Therefore,

$$\widehat{\Delta}(z; \mathcal{P}, N, M) \mid H_0 \Rightarrow \mathrm{Logistic}\left(\tau \log\left(\frac{N_1}{N_0}\right), \tau\right). \tag{15}$$

In practice one may either standardize the input series to enforce $\tau = 1$, or use the plug-in estimate $\hat{\tau} = s^2$, yielding $\mathrm{Logistic}(\hat{\tau} \log(N_1/N_0), \hat{\tau})$ for critical values.

In applied settings, we recommend the following procedure for testing. First, if the series is nonstationary (trending), apply first differences: $\delta z_t = z_t - z_{t-1}$. If the differenced series exhibits non-seasonal autocorrelation, fit a non-seasonal ARMA model by BIC and extract the residuals $\hat{e}_t$. Finally, compute $\widehat{\Delta}(\hat{e}; \mathcal{P}, N, M)$ and compare to the appropriate quantiles of the null distribution. Under this workflow, the Logistic null distribution applies to $\hat{e}_t$ with $\tau = \hat{\sigma}_e^2$. In our empirical experience, simple first-differencing provides adequate size control for most economic time series, obviating the need for formal ARMA pre-whitening. Section 5 provides simulation evidence supporting this claim under a range of colored-noise data-generating processes.

4. A Way Forward: Seasonal Adjustment

In Section 3, we formalized seasonality in the frequency domain as a property of the *relative peak dominance* over an *ex ante* partition of the frequency domain. We argued that, in the absence of seasonality, the spectral maxima in designated seasonal bins, $\mathcal{M}_1$, should be statistically indistinguishable from those in nonseasonal bins, $\mathcal{M}_0$. This definition relies on the relative height of spectral peaks, analogous to comparing the elevations of two mountain ranges relative to a common sea level. Under the null hypothesis of no seasonality, the difference in these maxima converges to a Logistic distribution with known parameters, providing a tractable test statistic.

In this section, we construct an adjustment operator that is the logical conjugate of this definition. Our objective is to remove the seasonal component in a manner that is (i) statistically robust, (ii) analytically transparent, (iii) agnostic to functional form, and (iv) scalable for production environments. To achieve this, we propose *Stochastic Spectral Imputation* (SSI), a non-parametric procedure that manipulates the series directly in the frequency domain. The intuition for SSI rests on the counterfactual implied by our definition. For a stationary process, if the excess power at seasonal frequencies were removed, the residual spectral density should resemble the underlying nonseasonal process.

[Insert Figure 1 Here]

For example, suppose we generate a series that is an additive function of trend, seasonal and irregular components: $Z = T + S + I$. We observe the series z_t with periodogram in Figure 1a,

and this series comes with a non-seasonal process with periodogram outlined in Figure 1b. In the real world, we don't know what the non-seasonal process is and as such Figure 1b is hidden. However, our objective is to modify the ordinates in the seasonal bins from Figure 1a so that the result is similar to that in Figure 1b overall. To accomplish this, SSI uses ordinates in the non-seasonal bins as donors to produce a new series with a periodogram that resembles that of Figure 1b. The remainder of this section will detail the SSI algorithm along with any options one may use in its configuration.

4.1. The SSI Algorithm

The Stochastic Spectral Imputation (SSI) algorithm operationalizes our definition of seasonality by treating the adjustment process as a surgical signal-extraction task in the frequency domain. Unlike filter-based methods that aggregate information across the spectrum, SSI targets only the specific frequency bands identified as seasonal during the detection phase thereby enforcing epistemic closure, a core contribution of this paper.

Let x_t be the observed series for $t = 1, \ldots, T$. We compute the discrete Fourier transform (DFT) coefficients, $d_x(\omega_j)$, such that

$$x_t = \frac{1}{T} \sum_{j=0}^{T_1} d_x(\omega_j) e^{i\omega_j t} \tag{16}$$

where $\omega_j = 2\pi jT$ for $j = 0, \ldots, T-1$. Using partitions $\mathcal{M}_1$ and $\mathcal{M}_0$ established in Section 3, we define the set of seasonal Fourier indices, $\mathcal{J}_S$, as those falling within the seasonal bins, and the donor set, $\mathcal{J}_D$, as the indices corresponding to nonseasonal background frequencies.

The procedure for constructing the adjusted series x_t^* consists of four primary steps. First, we identify $\mathcal{J}_S = \{j \in \mathcal{J} : \omega_j \in B_m$ for some $m \in \mathcal{M}_1$. To ensure the donor set represents the background noise floor, we optionally filter $\mathcal{J}_D$ using a quantile threshold, q_D, to exclude any non-seasonal frequency outliers that might distort the background distribution. The second step is the imputation; for each seasonal index $j \in \mathcal{J}_S$ we replace the observed magnitude, $|d_x(\omega_j)|$ with a draw from the empirical magnitude distribution of the donor set. Specifically, we sample k uniformly from $\mathcal{J}_D$ and set $|d^*(\omega_j)| = |d_x(\omega_k)|$. This means that the power in the seasonal bins is reduced to a level that is statistically indistinguishable from the nonseasonal background.

Step three involves the phase component of the DFT coefficient. Because the seasonal component is removed, the original phase information at seasonal frequencies is often no longer meaningful. SSI allows for phase ϕ_j to be kept from the original series, inherited from the donor, or drawn from $U(-\pi, \pi)$. For the Nyquist frequency and the DC component, we enforce the necessary real-value constraints. Finally, to ensure the adjusted series x_t^* remains real-values, step four enforces conjugate symmetry such that $d^*(\omega_{T-j}) = \overline{d^*(\omega_j)}$ for all j. The final adjusted series is obtained via the inverse discrete Fourier transform (IDFT) of the modified coefficients. A pseudo-code representation of SSI is covered in Algorithm 1.

[Insert Algorithm 1 Here]

4.2. Intuition for SSI

In Figure 2 we continue our example from Figure 1 to help build intuition and illustrate the impact of the SSI adjustment process. SSI moves iteratively through ordinates found to be within seasonal bins. These are the targets, and in Figure 2a we have marked one such target with a red box. A random ordinate is chosen from the donor pool with magnitude less than q_D, given by the dashed line in Figures 2a and 2b. The magnitude of the DFT coefficient corresponding to that ordinate is used in place of the magnitude from the target ordinate as illustrated by the open circle. The goal is that this should be in a location similar where the power of that frequency would be in the periodogram of the non-seasonal components. As mentioned earlier, the algorithm allows for the choice of the phase for that DFT coefficient to be kept from the target coefficient, its donor, or drawn randomly. Finally, conjugate symmetry is enforced for the second DFT coefficient implied by the target. This process is repeated for all ordinates inside $\mathcal{M}_1$. In Figure 2b we show the impact of a complete run through the algorithm. The open triangles represent the donated magnitudes that replaced ordinates in the seasonal bins. The final result of the SSI algorithm is the ordinates in Figure 2b, both in seasonal and non-seasonal bins, appear to be very similar to the oracle periodogram from Figure 1b. Moreover, one can visually see that there is no dominant peak in the seasonal bins compared to the non-seasonal bins. The value for $\widehat{\Delta}$ from Figure 1a is 9.48 with a p-value of 0.001 while the adjusted data given in Figure 2b produces $\widehat{\Delta} = -2.79$ with a p-value of 0.921.

[Insert Figure 2 Here]

The development of SSI was guided by four essential requirements for modern seasonal adjustment: (i) statistical robustness, (ii) analytical transparency, (iii) agnosticism toward functional form, and (iv) computational scalability. By utilizing the empirical distribution of the non-seasonal bins as a donor pool, SSI functions as a frequency-domain resampling procedure analogous to the bootstrap. This approach ensures the method remains computationally efficient and avoids the parametric commitments inherent in standard model-based pipelines. While this paper establishes the asymptotic justification for SSI, the framework's flexibility allows for the construction of uncertainty intervals through repeated draws, providing a path for future extensions into fully probabilistic seasonal adjustment. In the next section our simulations will demonstrate that SSI is not merely an alternative procedure, but the logical conjugate to our frequency-domain definition.

5. Monte Carlo Simulations

This section conducts simulation studies for our proposed testing and adjustment procedures. First, we show that the empirical null distribution of our test statistic almost exactly approximates its theoretical counterpart. Then, we contrast the power of our test statistic with popular alternatives proposed by the literature and find that it achieves comparable performance across several different specifications. Last, we explore the statistical properties of SSI under a known data generating process (DGP) and illustrate its advantages relative to X-13. We first describe the setup of and assumptions used in the DGP, which follows Busetti and Harvey (2003).

5.1. Setup

The data generating process is the basic structural model used in Busetti and Harvey (2003) consisting of the four equations below.

$$z_t = \mu_t + S_t'\gamma_t + \epsilon_t, \qquad \epsilon_t \sim \text{NID}(0, \sigma_\epsilon^2) \tag{17}$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \qquad \eta_t \sim \text{NID}(0, \sigma_\eta^2) \tag{18}$$
$$\beta_t = \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim \text{NID}(0, \sigma_\zeta^2) \tag{19}$$
$$\gamma_t = \boldsymbol{\alpha} + \rho\gamma_{t-1} + \kappa_t, \qquad \kappa_t \sim \text{NID}(0, \sigma_\kappa^2). \tag{20}$$

Here S_t is an $(N-1)$ vector of trigonometric seasonal variables and γ_t is an $(N-1)$ vector that controls their amplitudes. It is assumed that γ_t evolves according to an AR(1) process with some intercept $\boldsymbol{\alpha}$, an $(N-1)$ vector of constant values, and white-noise innovations κ_t. In Busetti and Harvey (2003) the value of ρ was fixed at one, meaning that the amplitudes evolved according to a random walk and, over a long sample period, the variance of the seasonal component would come to dominate that of z_t. By allowing γ_t to evolve as an AR(1) process we preserve the random walk aspect but also allow for a stationary, stochastically evolving set of amplitudes, and a fixed or memoryless evolving set of amplitudes if we so desire. We fix $T = 200$ across all simulations and report rejection rates consistent with a 5% significance level.

5.2. Evaluating the Empirical Null Distribution of $\widehat{\Delta}$

Before proceeding, we first evaluate the empirical null distribution as compared to the theoretical null of a $\text{Logistic}(\tau \log(N_1/N_0), \tau)$ derived earlier.[11] To do this we set $\sigma_\eta^2 = \sigma_\zeta^2 = \sigma_\kappa^2 = 0$ and $\boldsymbol{\alpha} = 0$ thereby allowing us to generate sample data from a white-noise process. We also set $M = 2N - 1$ leading to, in the monthly case, about two-thirds of the periodogram ordinates occurring in non-seasonal bins; this assumption assures convergence under the EVT conditions. Figure 3 plots the proposed test statistic under these conditions over the course of $50,000$ iterations. In Figure 3a, we can see that the histogram follows the theoretical density almost exactly and that the empirical 5% critical-value is nearly identical to the nominal value, producing size close to the nominal 5%. Figure 3b shows the null for quarterly observations, where we set $N = 4$ and $M = 7$. Again, we see that the empirical distribution mirrors the theoretical construct outlined in Section 3. Therefore, in finite samples the empirical null distribution converges to the nominal null distribution.[12]

[Insert Figure 3 Here]

5.3. Evaluating Empirical Power

Now, we explore the empirical power of our test statistic $\widehat{\Delta}$. For comparison, we calculate rejection rates for a battery of other popular test statistics found thoughout the literature: the Canova-Hansen (CH) test (Canova and Hansen, 1995), both the parametric and non-parametric (spectral) versions of the Busetti-Harvey test (BHp and BHs) (Busetti and Harvey, 2003), and the McElroy-Roy visual significance test (VS) McElroy and Roy (2021). Following Busetti and Harvey

[11] For simplicity in evaluation we scale the series such that $\tau = 1$ as indicated thus removing variation due to sample variance entering the empirical null distribution.

[12] However, more work is needed to understand the implications of the choice of M under finite T.

(2003) we set the lag parameter, m, for appropriate test statistics to a full year of observations. Following McElroy and Roy (2021) we set $\delta = 30$, a similar parameter to our M, and retain other suggested parameters. To make the VS test null comparable to that of $\widehat{\Delta}$ we jointly test the significance of all peaks in the spectrum using the Benjamini-Hochberg procedure to control the family-wise error rate (Benjamini and Hochberg, 1995).[13] For the other tests, a joint test statistic, null distribution, and critical values are all provided and we use them accordingly.

We examine empirical rejection rates for these tests under three different assumptions on the behavior of γ_t, the process that controls the amplitudes of the seasonal variables S_t.

1. γ_t evolves as a random walk ($\rho = 1$)
2. γ_t is a stationary process with high persistence ($\rho = 0.95$), where we fix $\boldsymbol{\alpha} = 0.01$
3. γ_t is a memoryless process ($\rho = 0$), where we vary $\boldsymbol{\alpha} \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$

In all three cases we hold $\sigma^2_\epsilon = 1$ and $\sigma^2_\zeta = 0$ constant while varying $\sigma^2_\eta \in \{0.0, 0.5\}$ and $\sigma^2_\kappa \in \{.01, .02, .03, .04, .05\}$. For cases where $\sigma^2_\eta > 0$, we assume that the data are non-stationary and take first differences, rather than first conducting a test for stationarity. Looking across this set of parameters shows how empirical power changes as the share of the variance of z_t that is driven by $S_t'\gamma_t$ grows. Moreover, these scenarios capture conditions commonly encountered in applied work using seasonal data, enabling us to evaluate performance in realistic settings.

Random Walk Amplitudes Table 2 shows the empirical rejection rates across tests and parameter constellations for the first scenario. Across columns, we report rejection rates under a 5% significance level for: the test statistic we introduce in this paper ($\widehat{\Delta}$), the non-parametric spectral Busetti-Harvey test (BHs), the parametric Busetti-Harvey test (BHp), the Canova-Hansen test (CH), the Quartically Seasonal test (QS), and the McElroy-Roy visual significance test (VS). In the final column, we report the proportion of total variance accounted for by the seasonal component S_t ($\%Var(S)$).

[Insert Table 2 Here]

Concentrating first on the case where $\sigma^2_\eta = 0$ (in other words, when the trend term in the simulated process z_t is purely deterministic) and monthly simulated data, we can see that the power of all tests increases as the variance σ^2_κ of the innovations to the scaling process γ_t grows larger. Intuitively, this makes sense as the seasonal component of the process must comprise a large enough fraction of the total variance in order for seasonality tests to detect it. However, we see that this threshold is quite low for $\widehat{\Delta}$: once $\%Var(S)$ reaches about 25%, our test already achieves a power of over 70%. Moreover, $\widehat{\Delta}$ strongly dominates all but the VS and QS tests across values of σ_κ; its power is in line with these two test statistics for $\sigma_\kappa \geq 0.03$.

The next panel ($\sigma^2_\eta = 0$) shows that we obtain similar results if we allow the trend term to stochastically drift.[14] We obtain similar results whereby $\widehat{\Delta}$ outperforms most of the other tests and is comparable to VS and QS at higher levels of σ_κ.

[13] A big thank you to Tucker McElroy who graciously provided us replication code for this test statistic in R with the recommended tuning parameters and suggested the Hochberg procedure to correct for multiple comparisons.

[14] Note that this increases (decreases) the variance of the irregular (seasonal) component by assumption.

For quarterly observations (bottom two panels) we obtain largely similar results. When $\sigma_\eta^2 = 0$, we find that $\widehat{\Delta}$ achieves comparable power to the other tests under consideration, though both QS and VS exhibit greater performance. When $\sigma_\eta^2 = 0.5$, $\widehat{\Delta}$ strictly dominates the other tests under consideration, including both QS and VS. We conclude that our test has good power under a process that displays nonstationary seasonality.

Stationary but Persistent Amplitudes Next, we explore the power of our test when the process that scales the amplitude of the seasonal component is stationary, yet highly persistent ($\rho = 0.95$). Table 3 displays our results for the same set of parameter values as in the previous section.

[Insert Table 3 Here]

For the case of monthly simulated data, we can see that $\widehat{\Delta}$ has reasonably good power relative to the other tests. The exception is the QS test, which strongly outperforms all others in this configuration. The reason is that the parameter values in this case induce large, positive autocorrelations in the simulated process at seasonal lags, which the QS test (a variation of the Ljung-Box test) is specifically designed to detect.

For quarterly observations with $\sigma_\eta^2 = 0$, the power of our test statistic is quite low, but is dominates the other tests under consideration besides VS and QS. With $\sigma_\eta^2 = 0.5$, $\widehat{\Delta}$ dominates all other test statistics under consideration in terms of power. We conclude that $\widehat{\Delta}$ is reasonably well-suited to detect this form of "stationary seasonality" as compared with other popular alternatives.

Memoryless Amplitudes Lastly, we consider a scenario with no persistence in the amplitude scaling process γ_t, in other words when $\rho = 0$. In this case, we vary the mean of γ_t between 0.1 and 0.5 while keeping the variance of its innovations constant. Table 4 displays the results.

[Insert Table 4 Here]

Again, our test statistic displays good power as compared with the other tests we consider. In most cases, $\widehat{\Delta}$ achieves performance that is either in line with or dominates the other tests. As the relative variance of S_t increases, the power of our test approaches 1 very quickly, similar to both the QS and VS tests. Collectively, we conclude from these exercises that the test statistics $\widehat{\Delta}$ we propose in this paper obtains good power under a wide range of scenarios, in line with the best performing alternatives in the literature.

5.4. Evaluating Adjustment with SSI

Having evaluated the empirical rejection rates of the test statistic developed in Section 3 we turn now to evaluating the efficacy of the Stochastic Spectral Imputation algorithm relative to the X-13 benchmark. As in Section 5.1, we employ a basic structural model with one critical modification: for all simulations, the variances of the trend components (σ_η^2 and σ_ζ^2) are set to zero. This produces a process z_t consisting solely of a seasonal component and an irregular component.

This constraint is central to our identification strategy. Since the true trend is deterministically zero, any trend component extracted by an adjustment procedure represents a "decomposition error"—the misidentification of low-frequency irregular variance as trend. This allows us to penalize "hallucinated" trends and treat the adjustment problem as a pure signal extraction task where the target is the precise recovery of the white-noise irregular component, ϵ_t. We assume for these

simulations that the presence of seasonality has already been established.

Judging the quality of a seasonal adjustment is a multidimensional problem that we evaluate through a combination of time- and frequency-domain metrics. The first is the Variance Ratio (VR), defined as $\mathrm{Var}(\hat{e}_t)/\mathrm{Var}(\epsilon_t)$. This serves as a detector for "over-smoothing." A ratio significantly below unity indicates that the adjustment process has suppressed variance in the irregular component, effectively discarding signal along with noise. As noted by Maravall (1993), standard linear filters often reduce the variance of the adjusted series below that of the true non-seasonal component.

Related to variance, we evaluate the Mean Absolute Error (MAE) of $\hat{e}_t$ relative to ϵ_t. We prioritize MAE over Mean Squared Error (MSE) because linear filters (like X-13) mechanically alter every ordinate in the frequency domain, whereas the SSI algorithm is surgical, leaving non-seasonal ordinates largely untouched. MSE tends to be dominated by large errors, whereas MAE provides a clearer aggregate measure of these small, broadband spectral changes. Finally, in the time domain, we evaluate the correlation between $\hat{e}_t$ and ϵ_t. While VR and MAE measure scale and distance, correlation measures dynamics; a high correlation indicates that the adjustment procedure has correctly preserved the phase and timing of structural innovations.

In the frequency-domain we will also measure the Spectral MAE,the mean absolute difference between periodogram ordinates of the adjusted data, $I_{\hat{e}}(\omega)$, and those of the true irregular, $I_{\epsilon}(\omega)$. Since the target irregular is white noise, this metric captures the global spectral distortion introduced by the adjustment process, with an optimal outcome of zero. Additionally, linear filters used in X-13 are known to create spectral dips or troughs at seasonal frequencies (McElroy and Roy, 2017). To quantify this, we calculate a Spectral Dip statistic: the average absolute difference between $I_{\hat{e}}(\omega)$ and $I_{\epsilon}(\omega)$ specifically within the designated seasonal frequency bins, $\mathcal{M}_1$. A smaller value indicates that the adjustment method successfully removed the seasonal peaks without carving out artificial valleys in the spectrum.

[Insert Table 5 Here]

Table 5 provides the results across both monthly and quarterly based series with varying levels of seasonal signal-to-noise. Across all specifications, X-13 consistently produces a Variance Ratio significantly below unity (e.g., ≈ 0.75 in monthly panels and ≈ 0.50 in quarterly panels). This confirms the well-documented tendency of linear filters to over-smooth the data by suppressing variance across the spectrum, creating a downward bias in the variance of the adjusted series. In contrast, SSI maintains a variance ratio near 1.0 across almost all specifications.

This over-smoothing behavior is supported by the spectral diagnostics, specifically the spectral bias or dip statistic. Our results show that X-13 systematically introduces deeper troughs at seasonal frequencies consistent with findings of McElroy and Roy (2017) and others. Overall, X-13 produces periodogram ordinates that are on average 50% lower than those produced by SSI relative to the ordinates from the true irregular term. By design, the filters used by X-13 are aggressive at dampening the power at seasonal frequencies in an effort to remove the seasonal signal, perhaps too aggressive given the spectral troughs it creates. Meanwhile, SSI avoids this by imputing power from a non-seasonal donor set rather than mechanically suppressing it. Because it is the logical extension of the definition of seasonality provided in Section 3 it is a local approach to removing that power. Consequently, there is less of a trough created by the process and, as

shown through the Spectral MAE in concert with the dip measure, SSI preserves the spectral density of the irregular component in a manner demonstrably better than the benchmark.

In the time-domain, a notable trade-off emerges between the smoothness of the series and structural fidelity. X-13 consistently achieves a lower MAE due to its variance-reducing properties. However, this smoothness comes at the cost of accuracy. SSI consistently achieves a higher correlation with the true irregular component, suggesting that while X-13 produces a "cleaner" looking series, SSI more accurately tracks the true irregular component. Thus, for users prioritizing the recovery of a true generating process (or an approximation thereof) over visual smoothness, SSI offers a robust alternative that mitigates the spectral distortions inherent in filter-based adjustments.

6. Empirical Examples

To illustrate the performance of our adjustment procedure relative to the current practice, we compare the results of using SSI versus X-13 on popular economic time series. We download both non-seasonally adjusted and seasonally adjusted published series from the Bureau of Labor Statistics (BLS) and Census Bureau and examine the characteristics of the published non-seasonally adjusted data, the published X-13 seasonally adjusted data, and the SSI-adjusted data.

6.1. Housing Starts

First, we consider new residential construction (aka "housing starts"), which are published monthly by the Census Bureau jointly with the Department of Housing and Urban Development (HUD). Housing starts measure the number of privately owned housing units on which construction begins and serve as a key indicator of housing market and broader economic conditions. They are closely watched by home builders, economists, policymakers, and investors alike because they reflect current housing demand and help forecast future construction employment and home sales.[15]

[Insert Figure 4 Here]

Figure 4 shows the resulting series from our adjustment procedure as well as some common diagnostic tests for seasonality. In the top-left panel, we plot the nonseasonally adjusted and seasonally adjusted data for the full period (1990–present). Two observations stand out. First, SSI results in a series that tracks the published, X-13 adjusted series very closely. We view this as a strength of SSI: it does not produce seasonally adjusted data that differ wildly from published aggregates, yet its transparency and known statistical properties make it a desirable alternative. Second, SSI results in a series that is slightly more volatile than X-13. We also view this property as a strength of SSI: it does not *over-smooth* the underlying data, which is a concern when using filter-based methods such as X-13.[16]

Reliably separating out seasonal patterns from the effects of underlying economic fundamentals is particularly important around business-cycle "turning points." Therefore, in the top-right panel, we zoom in on more recent data in order to examine the performance of SSI during the COVID

[15] Data are available at https://www.census.gov/econ/currentdata/ under the New Residential Construction heading. We download the Housing Units Started series.

[16] We elaborate more on this point below.

period. We can see from the figure that housing starts declined significantly throughout 2020, but seasonal factors were clearly at play. In particular, in the first quarter of 2020, the beginning of the COVID-19 pandemic coincided with the end of the winter housing market, which is typically less active compared to the rest of the year. Therefore, some of the observed softness in housing during this period was likely due to "true" seasonal factors. We can see from the figure that both seasonally adjusted series pick up on this pattern and suggest that the underlying level of housing starts was considerably higher (about 0.10 log points) than the unadjusted data during the first few months of 2020. However, the figure shows that X-13 overstates the level of housing starts during these months relative to SSI. Moreover, since housing starts under both SSI and X-13 dropped sharply in the remainder of 2020, X-13 overstates the contraction in the housing market. This feature of seasonal adjustment – accurately removing seasonal noise to isolate changes in economic fundamentals – is especially important because it allows data users to draw reliable inferences from published data.

The bottom two panels of Figure 4 dig into the mechanics of how each seasonal adjustment procedure operates. The bottom-left panel shows the autocorrelation function (ACF), a popular tool in time series analysis, for the three series: unadjusted, X-13 adjusted, and SSI adjusted. In the figure, we highlight the "seasonal lags" – those that occur around 12, 24, and 36 months – in order to concentrate on the relevant behavior. We can see that the autocorrelation of the unadjusted series with its seasonal lags is quite high; it is outside of the significance bands at these intervals. In other words, this diagnostic tool deems the series to be strongly seasonal, making it a useful test case. Next, we can see that SSI results in values of the ACF that occur inside the significance bands at all seasonal lags – SSI is successful at dampening these seasonal autocorrelations. Last, we can also see that X-13 significantly reduces the magnitudes of seasonal autocorrelations in most instances, with the exception of the value at $t - 12$. In this instance, the ACF value under the X-13 adjusted series is actually significantly negative. Therefore, the X-13 procedure took a series that is strongly positively correlated with its own value 12 months ago and produced a series that is strongly *negatively* correlated with its own value 12 months ago.

To further understand the behavior of X-13 relative to SSI, we examine its periodogram in the bottom-right panel of Figure 4. We normalize the y-axis of the periodogram using a log base 10 scale and again highlight the ordinates that occur in the seasonal bins. Unsurprisingly, the periodogram of the unadjusted series (black dots) has large spikes inside the seasonal bins.[17] For the periodogram of the SSI adjusted series, we can see that this is no longer the case: the values of the periodogram in the seasonal bins are akin to their nonseasonal counterparts. Again, X-13 displays similar behavior, but in some instances its periodogram ordinates fall below those of the other two series in the seasonal bins. Hence, X-13 negatively skews the height of the periodogram at seasonal frequencies, reminiscent of the over-smoothing behavior it displays in the time series.

6.2. Nonfarm Payrolls

Next, we repeat this exercise for nonfarm payroll employment published monthly by the BLS in the Employment Situation report, which is closely followed by both policymakers and financial market participants. This series measures the number of paid jobs in the economy excluding farm workers, private household employees, unpaid family workers, and the self-employed in agriculture,

[17] Our test statistic that is based on this logic has a p-value of 0 to machine precision.

covering roughly 80 percent of total employment in the U.S. economy.[18]

[Insert Figure 5 Here]

The top two panels again show that SSI produces results very similar to X-13 in the time series. Adjusted nonfarm payrolls under SSI are less volatile than the unadjusted series, as with X-13, and SSI does not materially alter the behavior of payroll employment during the COVID period, which was marked by historically large month-to-month changes that complicated seasonal adjustment considerably. That SSI can be implemented mechanically, without judgmental adjustments, highlights a key advantage of our approach.

Next, the bottom left and right panels show the ACF and periodogram, respectively. Again, we can see that for both X-13 and SSI, the adjusted series no longer has large, positive ACF values at 12, 24, and 36 months. If anything, SSI introduces an alternating pattern in the seasonal autocorrelations that was not necessarily present in the unadjusted data, as it operates using trigonometric functions, though these values are not significant.[19] By this metric, SSI is therefore successful in removing seasonal time-dependence in a statistical sense. The same conclusion holds in the frequency domain, where the figure shows that seasonal periodogram ordinates under the SSI adjusted series are again in line with their nonseasonal counterparts. In particular, the spikes in the periodogram at seasonal frequencies in the unadjusted series (black dots) disappear under SSI (red dots). Here, we again see evidence of over-smoothing by X-13: periodogram dispersion in the seasonal bins (blue dots) is much smaller than in the nonseasonal bins. This suggests that X-13's filtering methods suppress seasonal variability more than is actually warranted.

We conclude from these empirical examples that SSI is well-suited for seasonal adjustment at scale. The selected series are among the most closely followed economic indicators produced by statistical agencies. Our findings demonstrate that SSI not only produces results qualitatively similar to current practice but also yields adjusted time series with superior statistical properties, as illustrated by the ACF and periodogram.

7. Conclusion

This paper argues that the central difficulty in seasonal adjustment is conceptual rather than computational. In the absence of a unified definition of seasonality, diagnostics and adjustment procedures cannot be coherently aligned, resulting in heterogeneous treatment across producers and making clear communication between data providers and users difficult. We address this gap by introducing a non-tautological, frequency-domain definition based on relative peak dominance that yields a well-defined population object. From this object, we derive a test statistic with a known and analytically tractable null distribution and an adjustment operator — Stochastic Spectral Imputation (SSI) — that follows directly from the same logical construct. The result is an internally consistent structure in which definition, detection, and adjustment are mutually reinforcing rather than loosely connected.

[18] Data are available at data.bls.gov under the Current Employment Statistics - CES heading. We download the series IDs: CES0000000001 and CEU0000000001.

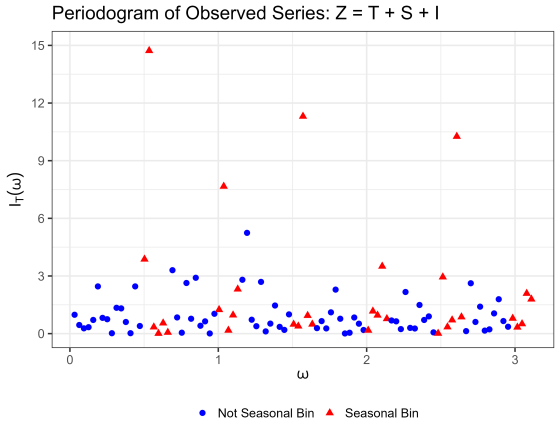
[19] Observe that in lags 11, 12, and 13, the ACF values under SSI are negative, positive, and negative. A similar pattern is seen in lags 23-25 and 35-37.

SSI operationalizes this framework by replacing ad hoc filtering with an adjustment operator that is the logical conjugate of the definition itself. Adjustment becomes a single, well-defined operation applied to a precisely stated object rather than the outcome of sequential diagnostic tuning. The resulting adjusted series closely track those produced by standard pipelines such as X-13, but the underlying structure is fundamentally different: seasonal adjustment becomes a well-posed statistical problem with a fixed target, a known sampling distribution, and an adjustment rule dictated by first principles rather than accumulated convention. The benefit of this approach is increased transparency for academics, data providers, policymakers, and practitioners alike.

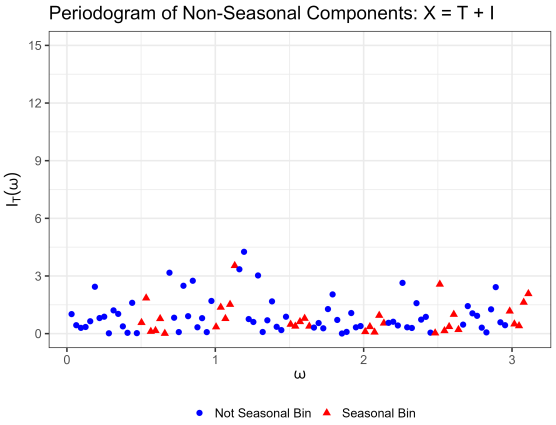
A. Tables and Figures

Figure 1. Conceptual Overview of Spectral Imputation

a. Observed Series Periodogram



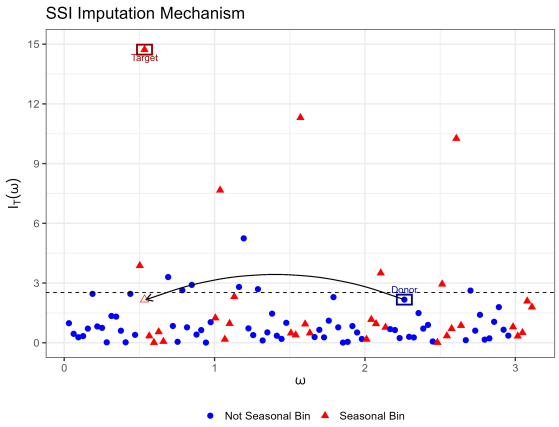
b. True Non-Seasonal Periodogram



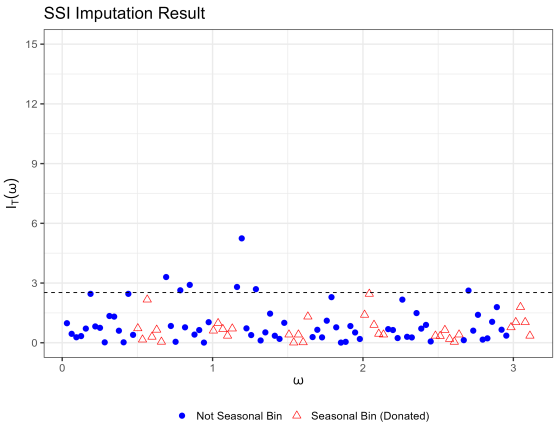
Note: Panel (a) shows the spectral density of a process with distinct seasonal peaks. Panel (b) shows the periodogram of the true non-seasonal components, something that is typically a hidden object.

Figure 2. Building Intuition for SSI

a. An Example of SSI



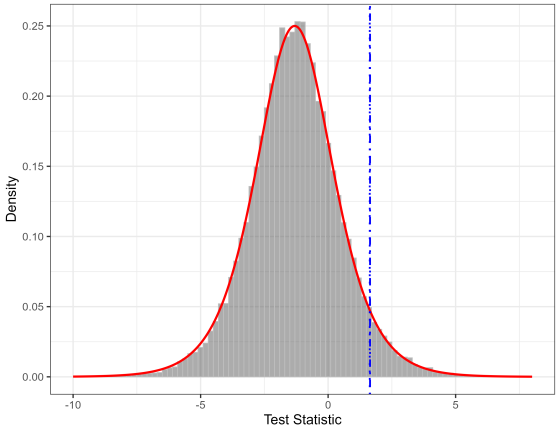
b. SSI Result



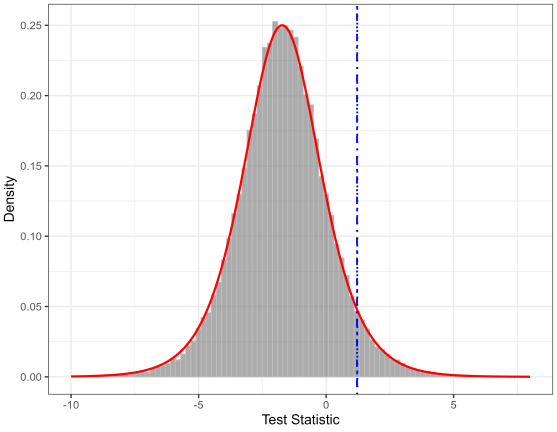
Note: Panel (a) shows a basic representation of the SSI algorithm through a single target-donor pair. Panel (b) shows the end result of the SSI algorithm, a periodogram that is comparable in structure to that found in Figure 1b.

Figure 3. A Comparison of the Empirical and Theoretical Null

a. Monthly Observations

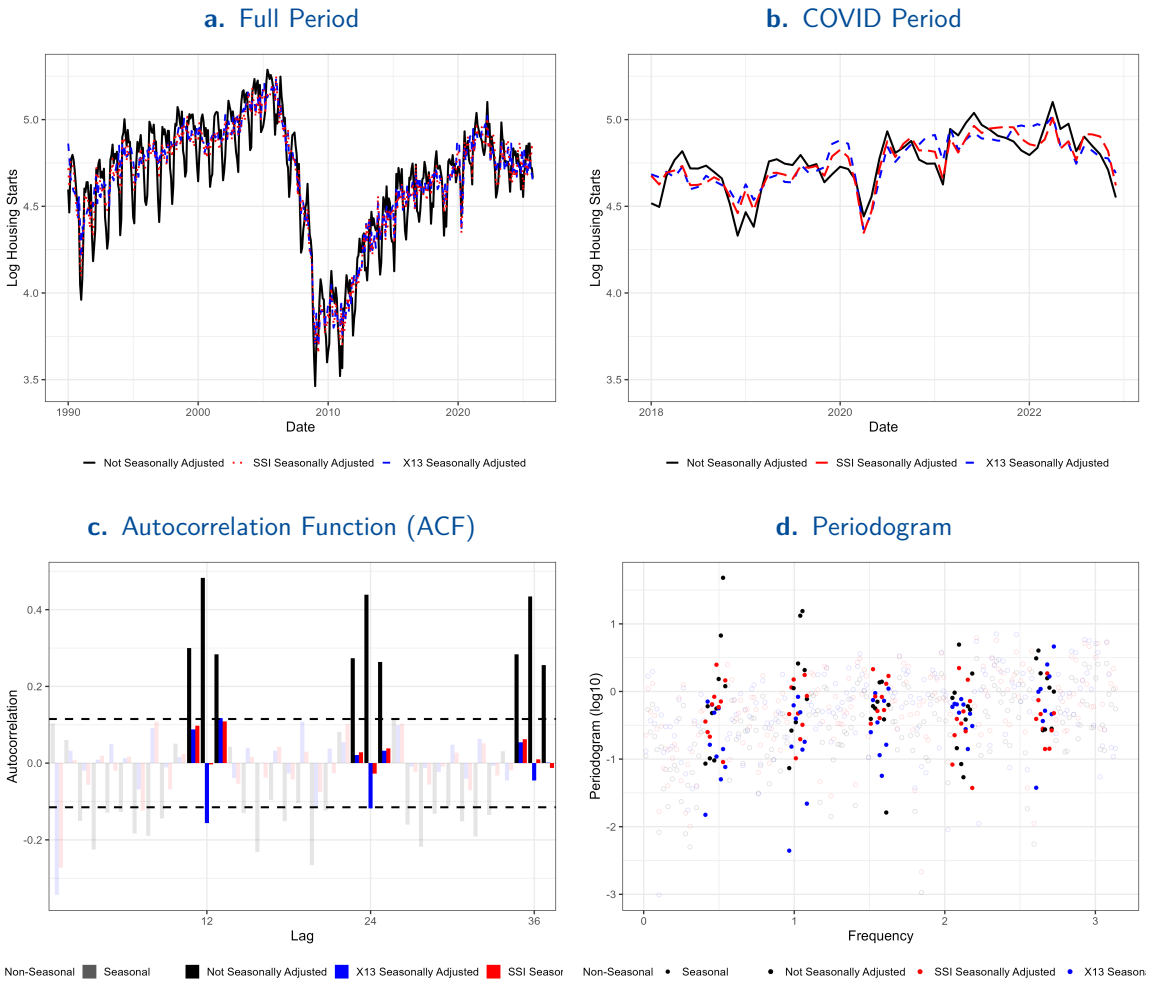


b. Quarterly Observations



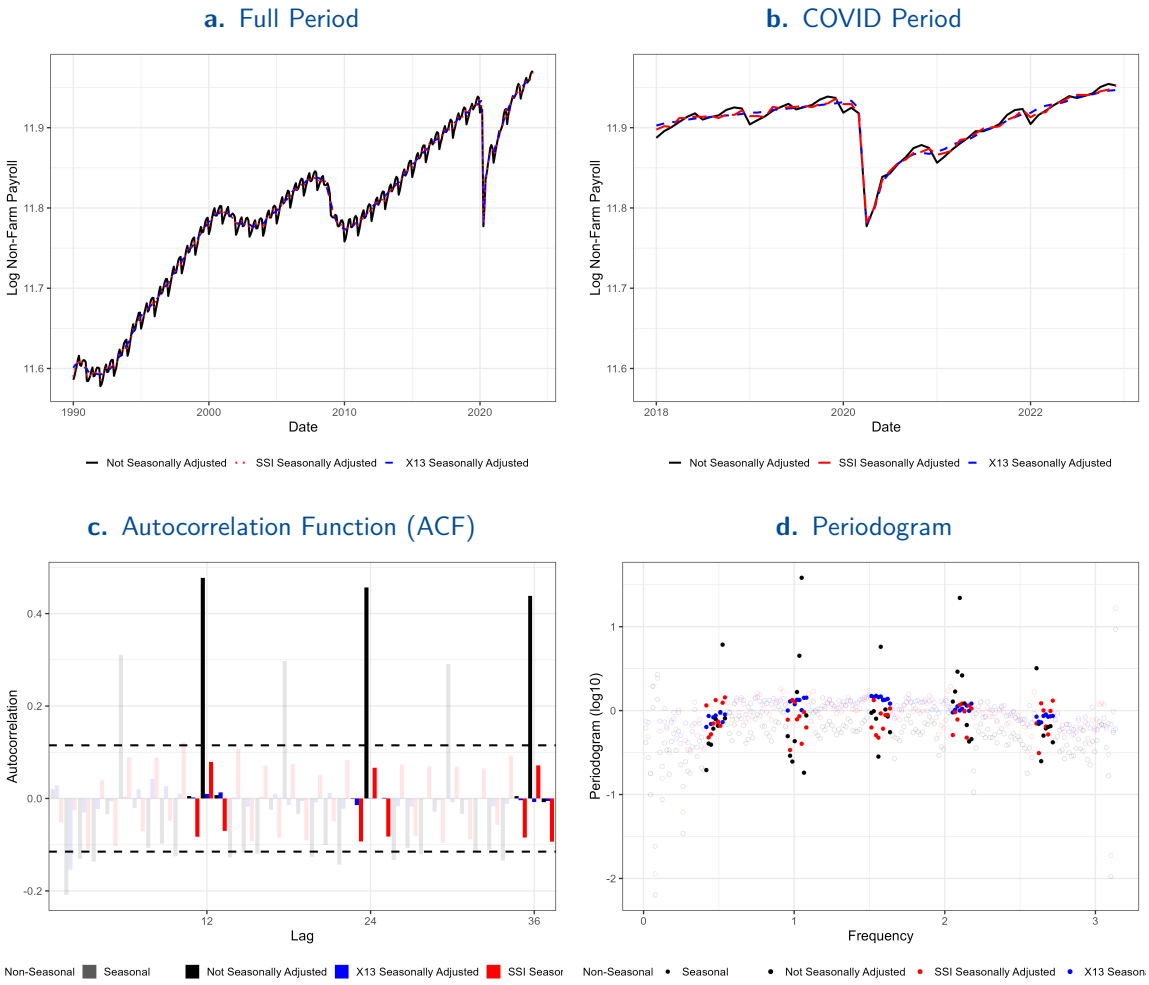
Note: We fix $T = 200$ across simulations and scale the series so that $\tau = 1$. Vertical lines are consistent with nominal and empirical 5% values.

Figure 4. Housing Starts



Note: Data are from the New Residential Construction report, issued jointly by the U.S. Census Bureau and the Department of Housing and Urban Development (HUD).

Figure 5. Nonfarm Payroll Employment



Note: Data are from the Employment Situation report produced by the U.S. Bureau of Labor Statistics (BLS) using data from the Current Employment Statistics (CES) program.

Table 1. Selected tests/diagnostics for seasonality (illustrative taxonomy)

Procedure	Domain	H_0	H_1
Friedman (1937)	Time	Stationary	Deterministic
Kruskal and Wallis (1952)	Time	Stationary	Deterministic
Priestley (1981)	Frequency	Stationary	Deterministic
Hylleberg et al. (1990)	Time	Unit Root	Stationary
Canova and Hansen (1995)	Time	Deterministic	Unit Root
Busetti and Harvey (2003)	Frequency	Zero	Deterministic or Unit Root
Lytras et al. (2007)	Time	Stationary	Deterministic
McElroy and Holan (2009)	Frequency	Zero	Stationary
Maravall (2012)	Time	Zero	Stationary
McElroy and Roy (2021)	Frequency	Mild Persistence	Strong Persistence
McElroy (2021)	Frequency	Strong Persistence	Mild Persistence

Note: This table is a modified version of that found in McElroy and Roy (2022).

Table 2. Stochastic Seasonality ($\rho = 1$)

Rates	σ_κ	$\widehat{\Delta}$	BHs	BHp	CH	QS	VS	$\%\mathrm{Var}(S)$
		Monthly Observations						
$\sigma_\eta^2/\sigma_\epsilon^2 = 0$	0.01	0.19	0.12	0.13	0.09	0.32	0.25	0.08
	0.02	0.72	0.40	0.42	0.26	0.88	0.75	0.24
	0.03	0.95	0.75	0.77	0.53	0.99	0.96	0.41
	0.04	0.99	0.92	0.93	0.76	1.00	0.99	0.55
	0.05	1.00	0.98	0.98	0.89	1.00	1.00	0.66
$\sigma_\eta^2/\sigma_\epsilon^2 = .5$	0.01	0.27	0.08	0.17	0.05	0.28	0.13	0.01
	0.02	0.68	0.28	0.43	0.16	0.79	0.53	0.03
	0.03	0.91	0.58	0.71	0.34	0.97	0.84	0.06
	0.04	0.97	0.81	0.88	0.55	1.00	0.96	0.10
	0.05	0.99	0.93	0.95	0.75	1.00	0.99	0.15
		Quarterly Observations						
$\sigma_\eta^2/\sigma_\epsilon^2 = 0$	0.01	0.07	0.09	0.08	0.08	0.11	0.15	0.02
	0.02	0.24	0.23	0.23	0.21	0.33	0.43	0.06
	0.03	0.50	0.43	0.44	0.41	0.64	0.68	0.13
	0.04	0.66	0.62	0.63	0.59	0.82	0.82	0.20
	0.05	0.79	0.74	0.76	0.72	0.92	0.88	0.28
$\sigma_\eta^2/\sigma_\epsilon^2 = .5$	0.01	0.23	0.09	0.18	0.09	0.12	0.09	0.00
	0.02	0.39	0.21	0.33	0.20	0.30	0.29	0.01
	0.03	0.61	0.38	0.50	0.36	0.54	0.54	0.01
	0.04	0.79	0.56	0.67	0.54	0.74	0.70	0.02
	0.05	0.88	0.70	0.79	0.68	0.86	0.81	0.04

Table 3. Stochastic Seasonality ($\rho = 0.95$)

Rates	σ_κ	$\widehat{\Delta}$	BHs	BHp	CH	QS	VS	$\%\mathrm{Var}(S)$
		Monthly Observations						
$\sigma_\eta^2/\sigma_\epsilon^2 = 0$	0.01	0.49	0.06	0.06	0.05	0.90	0.70	0.19
	0.02	0.51	0.07	0.06	0.06	0.89	0.70	0.20
	0.03	0.51	0.08	0.08	0.07	0.91	0.70	0.22
	0.04	0.54	0.11	0.13	0.09	0.91	0.70	0.24
	0.05	0.56	0.16	0.17	0.13	0.92	0.71	0.27
$\sigma_\eta^2/\sigma_\epsilon^2 = .5$	0.01	0.63	0.05	0.11	0.04	0.82	0.42	0.02
	0.02	0.62	0.05	0.12	0.04	0.81	0.43	0.02
	0.03	0.65	0.06	0.14	0.05	0.83	0.43	0.02
	0.04	0.64	0.09	0.18	0.06	0.83	0.43	0.03
	0.05	0.66	0.10	0.22	0.07	0.85	0.46	0.03
		Quarterly Observations						
$\sigma_\eta^2/\sigma_\epsilon^2 = 0$	0.01	0.24	0.06	0.06	0.06	0.46	0.63	0.07
	0.02	0.24	0.06	0.06	0.06	0.48	0.61	0.08
	0.03	0.26	0.08	0.08	0.08	0.51	0.62	0.08
	0.04	0.27	0.10	0.10	0.09	0.53	0.61	0.09
	0.05	0.30	0.13	0.14	0.12	0.58	0.61	0.10
$\sigma_\eta^2/\sigma_\epsilon^2 = .5$	0.01	0.49	0.07	0.13	0.06	0.43	0.42	0.01
	0.02	0.50	0.07	0.16	0.07	0.45	0.43	0.01
	0.03	0.50	0.08	0.17	0.08	0.46	0.42	0.01
	0.04	0.51	0.11	0.20	0.10	0.48	0.42	0.01
	0.05	0.54	0.13	0.23	0.12	0.52	0.42	0.01

Table 4. Deterministic Seasonality (Fixed $\rho = 0$)

Rates	γ_0	$\widehat{\Delta}$	BHs	BHp	CH	QS	VS	$\%\mathrm{Var}(S)$
		Monthly Observations						
$\sigma^2_\eta/\sigma^2_\epsilon = 0$	0.10	0.12	0.05	0.05	0.05	0.25	0.19	0.06
	0.20	0.48	0.05	0.05	0.05	0.90	0.70	0.19
	0.30	0.92	0.05	0.05	0.04	1.00	0.98	0.35
	0.40	1.00	0.05	0.05	0.04	1.00	1.00	0.49
	0.50	1.00	0.05	0.05	0.05	1.00	1.00	0.60
$\sigma^2_\eta/\sigma^2_\epsilon = .5$	0.10	0.24	0.05	0.10	0.03	0.25	0.10	0.01
	0.20	0.64	0.04	0.11	0.04	0.82	0.43	0.02
	0.30	0.94	0.04	0.10	0.03	1.00	0.86	0.04
	0.40	1.00	0.05	0.10	0.04	1.00	0.99	0.08
	0.50	1.00	0.04	0.11	0.04	1.00	1.00	0.12
		Quarterly Observations						
$\sigma^2_\eta/\sigma^2_\epsilon = 0$	0.10	0.07	0.05	0.05	0.05	0.12	0.19	0.02
	0.20	0.26	0.05	0.05	0.05	0.49	0.65	0.08
	0.30	0.73	0.05	0.05	0.05	0.90	0.94	0.15
	0.40	0.97	0.06	0.05	0.05	1.00	1.00	0.25
	0.50	1.00	0.06	0.06	0.05	1.00	1.00	0.34
$\sigma^2_\eta/\sigma^2_\epsilon = .5$	0.10	0.27	0.06	0.13	0.06	0.15	0.12	0.00
	0.20	0.50	0.06	0.13	0.06	0.45	0.44	0.01
	0.30	0.84	0.06	0.12	0.06	0.85	0.82	0.02
	0.40	0.98	0.06	0.13	0.05	0.99	0.98	0.03
	0.50	1.00	0.06	0.12	0.06	1.00	1.00	0.04

Table 5. Comparative Efficacy of SSI vs. X-13 Adjustment

	Monthly Observations										Quarterly Observations									
	Time Domain				Freq Domain				Diag		Time Domain				Freq Domain				Diag	
	MAE		Corr		MAE		Dip		VR		MAE		Corr		MAE		Dip		VR	
Param	SSI	X13	SSI	X13	SSI	X13	SSI	X13	SSI	X13	SSI	X13	SSI	X13	SSI	X13	SSI	X13	SSI	X13
Panel A: Seasonal Unit Root ($\rho = 1$)																				
$\sigma_\kappa = 0.01$	1.10	**1.02**	**0.93**	0.83	**0.31**	0.50	**0.22**	0.38	**0.94**	0.74	1.09	**0.95**	**0.94**	0.71	**0.32**	0.88	**0.19**	0.39	**0.93**	0.48
$\sigma_\kappa = 0.02$	1.10	**1.02**	**0.90**	0.82	**0.37**	0.52	**0.21**	0.39	**0.97**	0.74	1.09	**0.95**	**0.93**	0.71	**0.32**	0.88	**0.18**	0.40	**0.93**	0.48
$\sigma_\kappa = 0.03$	1.11	**1.02**	**0.86**	0.81	**0.43**	0.54	**0.21**	0.39	**1.02**	0.76	1.09	**0.95**	**0.93**	0.71	**0.33**	0.88	**0.20**	0.40	**0.93**	0.48
$\sigma_\kappa = 0.04$	1.13	**1.03**	**0.82**	0.79	**0.49**	0.57	**0.23**	0.40	**1.09**	0.77	1.09	**0.95**	**0.92**	0.71	**0.34**	0.88	**0.19**	0.41	**0.93**	0.48
$\sigma_\kappa = 0.05$	1.15	**1.03**	**0.79**	0.77	**0.55**	0.59	**0.23**	0.38	**1.17**	0.80	1.09	**0.95**	**0.91**	0.71	**0.35**	0.88	**0.19**	0.41	**0.94**	0.48
Panel B: Stochastic Stationary ($\rho = 0.95$)																				
$\sigma_\kappa = 0.01$	1.10	**1.02**	**0.91**	0.83	**0.34**	0.50	**0.21**	0.38	**0.95**	0.74	1.09	**0.95**	**0.94**	0.72	**0.32**	0.88	**0.19**	0.39	**0.93**	0.48
$\sigma_\kappa = 0.02$	1.11	**1.02**	**0.90**	0.83	**0.35**	0.51	**0.21**	0.38	**0.96**	0.75	1.09	**0.95**	**0.94**	0.71	**0.32**	0.88	**0.19**	0.39	**0.93**	0.48
$\sigma_\kappa = 0.03$	1.11	**1.03**	**0.90**	0.82	**0.38**	0.53	**0.21**	0.38	**0.98**	0.76	1.09	**0.95**	**0.93**	0.71	**0.33**	0.87	**0.19**	0.39	**0.93**	0.49
$\sigma_\kappa = 0.04$	1.11	**1.03**	**0.88**	0.81	**0.41**	0.54	**0.21**	0.37	**1.00**	0.78	1.09	**0.95**	**0.93**	0.71	**0.33**	0.88	**0.20**	0.40	**0.93**	0.49
$\sigma_\kappa = 0.05$	1.12	**1.04**	**0.87**	0.80	**0.45**	0.57	**0.21**	0.37	**1.03**	0.80	1.09	**0.95**	**0.93**	0.71	**0.34**	0.88	**0.19**	0.40	**0.93**	0.49
Panel C: Deterministic Seasonality ($\rho = 0$)																				
$\gamma_0 = 0.10$	1.10	**1.02**	**0.94**	0.84	**0.30**	0.49	**0.21**	0.38	**0.94**	0.74	1.09	**0.95**	**0.94**	0.72	**0.32**	0.87	**0.19**	0.39	**0.93**	0.48
$\gamma_0 = 0.20$	1.10	**1.02**	**0.91**	0.84	**0.33**	0.50	**0.21**	0.39	**0.95**	0.74	1.09	**0.95**	**0.94**	0.72	**0.32**	0.88	**0.19**	0.39	**0.93**	0.48
$\gamma_0 = 0.30$	1.11	**1.02**	**0.89**	0.84	**0.37**	0.50	**0.21**	0.39	**0.97**	0.74	1.09	**0.95**	**0.94**	0.72	**0.32**	0.87	**0.20**	0.40	**0.93**	0.48
$\gamma_0 = 0.40$	1.11	**1.02**	**0.86**	0.84	**0.41**	0.49	**0.22**	0.38	**1.00**	0.74	1.09	**0.95**	**0.94**	0.71	**0.32**	0.87	**0.20**	0.40	**0.92**	0.48
$\gamma_0 = 0.50$	1.12	**1.02**	0.84	0.84	**0.45**	0.49	**0.21**	0.38	**1.04**	0.74	1.09	**0.95**	**0.94**	0.72	**0.32**	0.88	**0.19**	0.40	**0.93**	0.48

Note: Bold values indicate better performance (Minimum MAE, Maximum Correlation, Maximum Seasonal Dip Power, or Variance Ratio closest to 1).

Algorithm 1 Generalized Stochastic Spectral Imputation (SSI)

Require: Observed series x_t, Seasonal params P, N, Bin param M, Options: Band, PhaseMode, q_{donor}

Compute DFT: $d_x \leftarrow \text{DFT}(x_t)$
Identify seasonal bins $\mathcal{M}_1$ and non-seasonal bins $\mathcal{M}_0$ based on $\Omega_G(P, N)$
Define initial sets: $\mathcal{J}_S^0 \leftarrow \{j : \omega_j \in \mathcal{M}_1\}$, $\mathcal{J}_D^0 \leftarrow \{j : \omega_j \in \mathcal{M}_0\}$
if Band = "between_seasonal_extremes" **then**
    Determine frequency range $[m_{\min}, m_{\max}]$ from $\mathcal{M}_1$
    Restrict $\mathcal{J}_S \leftarrow \{j \in \mathcal{J}_S^0 : \omega_j \in [B_{m_{\min}}, B_{m_{\max}}]\}$
    Restrict $\mathcal{J}_D^{\text{cand}} \leftarrow \{j \in \mathcal{J}_D^0 : \omega_j \in [B_{m_{\min}}, B_{m_{\max}}]\}$
else
    $\mathcal{J}_S \leftarrow \mathcal{J}_S^0$, $\mathcal{J}_D^{\text{cand}} \leftarrow \mathcal{J}_D^0$
end if
Filter Donors: $\tau \leftarrow \text{Quantile}(\{|d_x(\omega_k)|^2\}_{k \in \mathcal{J}_D^{\text{cand}}}, q_{\text{donor}})$
$\mathcal{J}_D \leftarrow \{k \in \mathcal{J}_D^{\text{cand}} : |d_x(\omega_k)|^2 \leq \tau\}$
Initialize adjusted coefficients: $d^* \leftarrow d_x$
for $j \in \mathcal{J}_S$ **do**
    Sample donor index $k \sim \text{Uniform}(\mathcal{J}_D)$
    Impute magnitude: $|d^*(\omega_j)| \leftarrow |d_x(\omega_k)|$
    if $\omega_j = \pi$ (Nyquist Frequency) **then**
        Determine sign s:
        if PhaseMode = "random" **then** $s \sim \text{Uniform}(\{-1, 1\})$
        else if PhaseMode = "keep" **then** $s \leftarrow \text{sgn}(\text{Re}(d_x(\omega_j)))$
        else if PhaseMode = "donor" **then** $s \leftarrow \text{sgn}(\text{Re}(d_x(\omega_k)))$
        end if
        $d^*(\omega_j) \leftarrow s \cdot |d^*(\omega_j)|$
    else
        Determine phase ϕ:
        if PhaseMode = "random" **then** $\phi \sim \text{Uniform}(-\pi, \pi)$
        else if PhaseMode = "keep" **then** $\phi \leftarrow \text{Arg}(d_x(\omega_j))$
        else if PhaseMode = "donor" **then** $\phi \leftarrow \text{Arg}(d_x(\omega_k))$
        end if
        $d^*(\omega_j) \leftarrow |d^*(\omega_j)|e^{i\phi}$
        Enforce symmetry: $d^*(\omega_{T-j}) \leftarrow \overline{d^*(\omega_j)}$
    end if
end for
return Adjusted series $x_t^* \leftarrow \text{Re}(\text{IDFT}(d^*))$

References

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological) 57*(1), 289–300.

Busetti, F. and A. Harvey (2003). Seasonality tests. *Journal of Business & Economic Statistics 21*(3), 420–436.

Canova, F. (1992). An alternative approach to modeling and forecasting seasonal time series. *Journal of Business & Economic Statistics 10*(1), 97–108.

Canova, F. and B. E. Hansen (1995). Are seasonal patterns constant over time? a test for seasonal stability. *Journal of Business & Economic Statistics 13*(3), 237–252.

Chen, J. and G. Cornwall (2021). The dark side of the moon: Searching for the other half of seasonality. Technical report, U.S. Bureau of Economic Analysis.

Commission, E. and Eurostat (2015). *ESS guidelines on seasonal adjustment – 2015 edition*. Publications Office.

Commission, E. and Eurostat (2024). *ESS guidelines on seasonal adjustment – 2024 edition*. Publications Office.

Dagum, E. B. (1980). *The X-II-ARIMA seasonal adjustment method*. Statistics Canada.

Doppelt, R. (2023). Should macroeconomists use seasonally adjusted time series? structural identification and bayesian estimation in seasonal vector autoregressions.

Falkner, H. D. (1924). The measurement of seasonal variation. *Journal of the American Statistical Association 19*(146), 167–179.

Findley, D. F., B. C. Monsell, W. R. Bell, M. C. Otto, and B.-C. Chen (1998). New capabilities and methods of the x-12-arima seasonal-adjustment program. *Journal of Business & Economic Statistics 16*(2), 127–152.

Franses, P. H. (1994). A multivariate approach to modeling univariate seasonal time series. *Journal of Econometrics 63*(1), 133–151.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association 32*(200), 675–701.

Ghysels, E. and D. R. Osborn (2001). *The econometric analysis of seasonal time series*. Cambridge University Press.

Ghysels, E., D. R. Osborn, and P. M. Rodrigues (2006). Forecasting seasonal time series. *Handbook of economic forecasting 1*, 659–711.

Gómez, V. and A. Maravall (2001). Seasonal adjustment and signal extraction in economic time series. *A course in time series analysis*, 202–247.

Granger, C. W. (1978). Seasonality: causation, interpretation, and implications. In *Seasonal analysis of economic time series*, pp. 33–56. NBER.

Hylleberg, S., R. F. Engle, C. W. Granger, and B. S. Yoo (1990). Seasonal integration and cointegration. *Journal of econometrics 44*(1-2), 215–238.

Jaditz, T. (1994). Seasonality: economic data and model estimation. *Monthly Lab. Rev. 117*, 17.

Kallek, S. (1978). An overview of the objectives and framework of seasonal adjustment. *Seasonal analysis of economic time series*, 3–32.

Kruskal, W. H. and W. A. Wallis (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association 47*(260), 583–621.

Kuznets, S. (1933). *Seasonal Variations in Industry and Trade*. General series. National Bureau of Economic Research.

Lytras, D. P., R. M. Feldpausch, and W. R. Bell (2007). Determining seasonality: a comparison of diagnostics from x-12-arima. *US Census Bureau*.

Maravall, A. (1993). Stochastic linear trends: Models and estimators. *Journal of Econometrics 56*(1-2), 5–37.

Maravall, A. (2011). Seasonality tests and automatic model identification in tramo-seats. *Bank of Spain: Madrid, Spain*.

Maravall, A. (2012). Update of seasonality tests and automatic model identification in tramo-seats. *Bank of Spain*.

McElroy, T. (2008). A modified model-based seasonal adjustment that reduces spectral troughs and negative seasonal correlation. *Statistics*, 12.

McElroy, T. (2021). A diagnostic for seasonality based upon polynomial roots of arma models. *Journal of Official Statistics 37*(2), 367–394.

McElroy, T. and S. Holan (2009). A nonparametric test for residual seasonality. *Survey Methodology 35*, 67–83.

McElroy, T. and A. Roy (2017). Detection of seasonality in the frequency domain. *Statistics 1*(2).

McElroy, T. and A. Roy (2021). Testing for adequacy of seasonal adjustment in the frequency domain. *Journal of Statistical Planning and Inference 211*, 241–255.

McElroy, T. and A. Roy (2022). A review of seasonal adjustment diagnostics. *International Statistical Review 90*(2), 259–284.

Nerlove, M. (1964). Spectral analysis of seasonal adjustment procedures. *Econometrica: Journal of the Econometric Society*, 241–286.

Politis, D. N. and T. S. McElroy (2019). *Time series: A first course with bootstrap starter*. Chapman and Hall/CRC.

Priestley, M. (1981). Spectral analysis and time series.

Sims, C. A. (1993). Rational expectations modeling with seasonally adjusted data. *Journal of Econometrics 55*(1–2), 9–19.

Soukup, R. J. and D. F. Findley (1999). On the spectrum diagnostics used by x-12-arima to indicate the presence of trading day effects after modeling or adjustment. In *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 144–149.

United Nations, European Commission, I. M. F. O. f. E. C., Development, and W. Bank. (2025). System of national accounts 2025. Technical report, United Nations.

Webel, K. and D. Ollech (2018). An overall seasonality test based on recursive feature elimination in conditional random forests. In *Proceedings of the 5th international conference on time series and forecasting*, pp. 20–31.