Nonparametric HAC Estimation for
Time Series Data with Missing Observations

Deepa Dhume Datta and Wenxin Du

Nonparametric HAC Estimation for Time Series Data with Missing Observations[*]

Deepa Dhume Datta[†]   Wenxin Du[‡]

November 5, 2012

Abstract

The Newey and West (1987) estimator has become the standard way to estimate a heteroskedasticity and autocorrelation consistent (HAC) covariance matrix, but it does not immediately apply to time series with missing observations. We demonstrate that the intuitive approach to estimate the true spectrum of the underlying process using only the observed data leads to incorrect inference. Instead, we propose two simple consistent HAC estimators for time series with missing data. First, we develop the Amplitude Modulated estimator by applying the Newey-West estimator and treating the missing observations as non-serially correlated. Secondly, we develop the Equal Spacing estimator by applying the Newey-West estimator to the series formed by treating the data as equally spaced. We show asymptotic consistency of both estimators for inference purposes and discuss finite sample variance and bias tradeoff. In Monte Carlo simulations, we demonstrate that the Equal Spacing estimator is preferred in most cases due to its lower bias, while the Amplitude Modulated estimator is preferred for small sample size and low autocorrelation due to its lower variance.

Keywords: heteroskedasticity, serial correlation, robust inference, missing data
JEL Classifications: C22, C12, C14

1 Introduction

While use of the Newey and West (1987) estimator and its Andrews (1991) implementation have become standard practice for heteroskedasticity and autocorrelation (HAC) robust inference, analogous methods for series with missing observations are far from standardized. When data are missing, the Newey-West formulas do not immediately apply, and the formula for calculating the lagged autocorrelations that are required terms in conventional HAC formulas must be adjusted.

Current practices for working with missing data include treating the missing observations as non-serially correlated, or imputing or ignoring the missing observations. To our knowledge, there has not been formal justification of HAC estimation for robust inference in these contexts, and the effect of employing these work-around methods on the resulting inferences is generally unexplored in applied work. In this paper, we provide formal justification for two methods of HAC estimation, and we compare these two methods to other existing methods. We demonstrate that treating the data as equally spaced is preferred under very general conditions, and that treating the missing observations as non-serially correlated may be preferred in instances with small sample sizes or low autocorrelation. In general, we find that our two newly formalized methods are preferred to imputation.

Especially when our aim is to adjust inferences for serial correlation, it seems counterintuitive that we can either treat the data as equally spaced or treat the missing observations as non-serially correlated, since these procedures require us to depart from the original time structure or autocorrelation structure of the data. However, we show that these procedures both provide consistent estimators of the long-run variance of the observed series with missing data. Though many have suggested that we infer the spectrum of the underlying data from the observed data using the *Parzen estimator*, we show that the Parzen estimator is not the correct object for inference testing. Rather, we show that our *Amplitude Modulated estimator* (which treats the missing as non-serially correlated) and *Equal Spacing estimator*

(which treats the observed as equally spaced) are extremely simple to implement and can be used to generate asymptotically valid inferences.

These insights are particularly valuable given the *ad hoc* approaches widely found in the applied literature. For example, researchers often use imputation procedures to fill in the missing data. Imputation seems to expand the set of observations used for analysis, or at least prevents us from dropping data when some covariates are unobserved. However, because imputed data series are often smoothed versions of the underlying unobserved series, they will often lead to asymptotically valid but extremely poor finite sample performance. Furthermore, researchers using these methods rarely adjust their inferences for this induced serial correlation.

Additional evidence of the confusion in the literature stems from the implementation of the Newey-West HAC estimator in the popular statistical package, Stata. The *newey2* command implements Newey-West to obtain the standard error of the coefficient in a regression using time series or panel data. When observations are missing, the option "force" can be applied. This option will apply our Equal Spacing estimator to time series data and apply our Amplitude Modulated estimator to panel data. However, it should be possible to implement either the Equal Spacing estimator or the Amplitude Modulated estimator for time series data. Furthermore, the program will not apply the Amplitude Modulated estimator at all when some lags are unobserved. This condition is likely an artifact of the literature on estimating the long-run variance of the underlying series, which develops estimators that generally require all lags to be observed (Clinger and Van Ness, 1976; Parzen, 1963). Yet for robust inference, we show that the Amplitude Modulated and Equal Spacing estimators do not require all the lags to be observed.

A primary goal of this paper is to help researchers select the correct estimation procedure in applied work. To that end, we follow the style of Petersen (2009), which provides guidance for selecting standard errors in finance panel data sets. We formally present the Amplitude Modulated and Equal Spacing estimators of the long-run variance of the observed series,

and we review their asymptotic properties. We contrast these estimators with the Parzen estimator of the long-run variance of the underlying series. After presenting these theoretical contributions, we offer intuition for why the estimators work and how they are related to each other. To generate guidance for choosing the correct estimator, we conduct Monte Carlo simulation results for various sample sizes, correlation structures, and fractions of observations that are missing. In addition to testing our estimators using randomly missing data, we demonstrate the applicability of these estimators to a deterministic cyclical missing structure, as with daily financial data, which usually cover 5 of 7 days of the week. Finally, we discuss an empirical application using recursive regressions for commodities futures returns to demonstrate how the choice of estimator can affect the conclusion of empirical tests.

As a preview, our results demonstrate that the Amplitude Modulated and Equal Spacing estimators are both consistent under random and deterministic missing structures. In finite samples, we find that for the same fixed bandwidth, the Equal Spacing estimator is generally less biased than the Amplitude Modulated estimator, but has larger variances. Consequently, the Equal Spacing estimator is preferred when autocorrelation is high, as the bias will dominate the mean squared error in these cases. Conversely, when autocorrelation is low, variance dominates the mean squared error, and the Amplitude Modulated estimator is preferred. The precise cutoff between these cases depends on the sample size, and whether automatic bandwidth selection procedures are implemented.[1]

The remainder of the paper is structured as follows. Section 1.1 discusses some examples of missing data problems for which imputation and equal spacing methods have been applied, and provides a brief review of some of the related econometrics literature. Section 2 provides an overview of our estimators by applying them in a simple setting with missing observations. Section 3 formally defines the estimators and discusses their asymptotic and finite sample

---

[1]To make it easier for researchers to apply these estimators, we have posted Matlab code for both estimators on our websites. We also have posted a basic simulation code that reports empirical rejection rates, size-adjusted power, bias, and variance for the Equal Spacing, Amplitude Modulated, Impuation, and (full sample) Newey-West estimators. Researchers can use the simulation code to evaluate the performance of the estimators under customized sample size, autocorrelation, and missing structure before choosing which estimator to implement.

properties. Section 4 presents the application of these estimators to inference in a regression setting with missing observations. Section 5 describes the Monte Carlo simulations based on these estimators and the results. Section 6 presents an empirical application of the estimators using data on commodities returns. Finally, Section 7 concludes.

1.1 Relation to the Literature

This paper extends the HAC covariance literature to applications with missing observations. In general, the most commonly used HAC covariance matrix is the one proposed by Newey and West (1987) and further developed by Andrews (1991). The Newey-West estimator equals a weighted sum of lagged autocovariance matrices, in which the weights are calculated using the Bartlett kernel. Newey and West (1987) show that this estimator is positive semi-definite and heteroskedasticity and autocorrelation consistent. Andrews (1991) and Newey and West (1994) investigate the finite sample properties of these estimators and propose data-dependent bandwidths. Though some papers have proposed variants of the estimators discussed in these seminal papers, the estimators applied in most of the current literature remain largely unchanged from their original form.

There have been earlier attempts to estimate HAC covariance matrices when some observations are missing. The Parzen (1963) paper on spectral analysis for data series with missing observations focuses on estimating the autocovariances of the underlying process in the presence of missing observations, based on a specific cyclical structure of missing data. We contribute to this literature by pointing out that for robust inference, we generally require an estimate of the long-run variance of the observed series rather than the underlying. We differentiate between the Amplitude Modulated estimator and the *Parzen estimator*. Following Parzen (1963), both of these estimators involve recasting the observed series as an amplitude modulated series in which the value of the underlying series is set to zero when observations are missing. The observed time structure of the data is respected, and the lagged autocovariances are estimated using only the lagged pairs which are fully

observed. We show that while the Amplitude Modulated estimator is a consistent estimator of the long-run variance of the observed series, the Parzen estimator is a consistent estimator of the long-run variance of the underlying series. Along with these theoretical results, we provide simulation results that demonstrate consistency of the t-statistic constructed using the Amplitude Modulated estimator.

We also argue to extend the set of missing data structures to which these estimators can be applied. Other researchers have attempted to apply Parzen's work to a variety of missing data structures, including the Bernoulli structure of randomly missing variables and more general cyclical patterns of missing observations (Scheinok, 1965; Bloomfield, 1970; Clinger and Van Ness, 1976; Dunsmuir and Robinson, 1981a,b). While the literature has moved beyond Parzen's original application, it still is focused on applications with randomly missing observations. Yet, many common applications of missing data techniques are for data that have a deterministic missing structure. Our theoretical results and simulation exercise demonstrate that as long the missing data structure satisfies our independence assumption, we can apply the Amplitude Modulated and Equal Spacing estimators to settings in which the pattern of missing observations is deterministic instead of random. This extends the set of possible applications to include business daily data, for example, in which weekends could be considered missing data.

More recently, Kim and Sun (2011) construct a HAC estimator for the two-dimensional case robust to spatial heteroskedasticity and autocorrelation. The focus of the paper is not on missing data, and they do not distinguish the difference between the spatial spectrum of the underlying versus the observed process. However, they discuss the applicability of their method to irregularly observed spatial data. Reducing the spatial HAC on the irregular lattice to one-dimensional time series produces an estimator very similar to our Amplitude Modulated estimator. We clarify the subtlety between the underlying and observed spectrum and develop the Amplitude Modulated estimator in the context of time series with missing data.

Setting the theory aside, many researchers use imputation techniques when faced with missing observations in practice. For example, one common use of imputation is to temporally disaggregate data to generate quarterly data from annual series, or monthly data from quarterly series. The Denton method of imputation smooths data when generating these series by minimizing first-differences or second-differences (Denton, 1971). Relatedly, the Chow-Lin method uses a related indicator series that can be used to interpolate, distribute, or extrapolate data (Chow and Lin, 1971, 1976). When this method is used, some properties of the indicator series, including serial correlation, will be transferred to the imputed series. Even the simplest method of imputing data by naive linear interpolation will induce autocorrelation in the imputed series. Studies based on Monte Carlo simulations suggest that even for reasonably large sample sizes, inference methods based on Newey-West HAC covariance estimators result in significant overrejection when the serial correlation is high (den Haan and Levin, 1997). In an imputed series, the induced high autocorrelation exacerbates this distortion. Yet, researchers using these methods rarely adjust their inferences for this induced serial correlation (for two such examples, see Eaton, Kortum, Neiman, and Romalis (2011) and Forni, Monteforte, and Sessa (2009)). We show in our Monte Carlo simulations and our empirical application that our estimators are simple alternatives that avoid the problems associated with imputation.

To avoid the complication of adjusting HAC estimators for the method and extent of imputation, some researchers simply ignore the missing observations. Formally, this method amounts to relabeling the time index and treating observations as though they are equally spaced in time. While this method has no previous formal justification to our knowledge, it has been widely applied. For example, observations of daily financial data are generally treated as equally spaced consecutive observations, irrespective of their actual spacing in time (examples include Acharya and Johnson (2007), Beber, Brandt, and Kavajecz (2009), Jorion and Zhang (2007), Pan and Singleton (2008), and Zhang, Zhou, and Zhu (2009)). Yet, for prices that are affected by developments in global markets with different business weeks or

national holidays, the lack of price data on weekends and holidays could be treated as missing observations. In this paper, we formalize this Equal Spacing estimator and demonstrate its asymptotic consistency and finite sample performance.

In light of the existing confusion in the literature on HAC estimation with missing data, we provide our Equal Spacing and Amplitude Modulated estimators as alternatives. We discuss the finite sample properties of these estimators and provide simulation results that offer insight into the bias and variance tradeoffs between the two estimators, so that practitioners can make an informed choice before applying either one.

2 A simple example

To fix ideas, in this section we introduce each of our estimators in the context of a simple example using three weeks of daily gasoline prices. We reserve for the next section a more detailed discussion of the asymptotic and finite sample properties of the estimators and the practical considerations involved in choosing among them. Suppose we have gasoline price data $\{z_t\}$ for the first three weeks of the month:

Mon	Tues	Weds	Thurs	Fri	Sat	Sun
z_1	z_2	z_3	z_4	z_5	z_6	z_7
z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}
z_{15}	z_{16}	z_{17}	z_{18}	z_{19}	z_{20}	z_{21}

For clarity of exposition, suppose these data have already been demeaned, so that we have $E(z_t) = 0$. To estimate the long-run variance of the series $\{z_t\}$, we can apply the standard Newey-West estimator:

$$\hat{\Omega}^{NW} = \hat{\gamma}(0) + 2 \sum_{j=1}^{m} w(j,m)\hat{\gamma}(j),$$

where the Bartlett kernel, $w(j,m) = 1 - [j/(m+1)]$ if $j \leq m$ and $w(j,m) = 0$ if $j > m$, is used to weight the sample autocorrelations at each lag j:

$$\hat{\gamma}(j) = \frac{1}{T} \sum_{t=j+1}^{T} z_{t-j} z_t,$$

In our example, we can estimate the first lagged autocorrelation for the gasoline price series as:

$$\hat{\gamma}(1) = \frac{1}{21}[z_1 z_2 + z_2 z_3 + \ldots + z_{20} z_{21}].$$

Similarly, we estimate the third lagged autocorrelation as:

$$\hat{\gamma}(3) = \frac{1}{21}[z_1 z_4 + z_2 z_5 + \ldots + z_{18} z_{21}].$$

Note that the denominator in both cases is the total number of observations, T, rather than the number of observed lags, $T - j$.

Now suppose we have only business daily data, with missing data on weekends:

Mon	Tues	Weds	Thurs	Fri	Sat	Sun
z_1	z_2	z_3	z_4	z_5	.	.
z_8	z_9	z_{10}	z_{11}	z_{12}	.	.
z_{15}	z_{16}	z_{17}	z_{18}	z_{19}	.	.

When some data points are missing, we have a few choices for how we estimate the lagged autocovariances, $\hat{\gamma}(j)$, that are components of the long-run variance, $\hat{\Omega}$. Especially in the context of business daily data, one very common procedure is to ignore the missing data. In this case, we would treat the observed prices as equally spaced in time. When estimating the first lagged autocovariance for our *Equal Spacing estimator*, we would treat the jump from Friday to Monday (e.g. day 5 to day 8) as a one day lag:

$$\hat{\gamma}^{ES}(1) = \frac{1}{15}[z_1 z_2 + z_2 z_3 + z_3 z_4 + z_4 z_5 + \underline{z_5 z_8} + z_8 z_9 + z_9 z_{10} + \ldots + z_{18} z_{19}].$$

Similarly, the third autocovariance would be estimated as:

$$\hat{\gamma}^{ES}(3) = \frac{1}{15}[z_1 z_4 + z_2 z_5 + \underline{z_3 z_8 + z_4 z_9 + z_5 z_{10}} + z_8 z_{11} + z_9 z_{12} + z_{10} z_{15} + \ldots + z_{16} z_{19}].$$

Since we have effectively reduced the sample size by ignoring the missing days, the denominator for each estimated lagged autocovariance is the total number of observed data points, or 15 in our example.

Alternatively, we could estimate the lagged autocovariances using only the instances in which we observe data with the right spacing. We call this the *Amplitude Modulated estimator*, because we are effectively modulating, or setting to 0, the value (or amplitude) of the series on the missing days. Using this estimator, to estimate the first lagged autocovariance in our example, we would use the lag from day 4 to day 5 and then skip to the lag from day 8 to day 9:

$$\hat{\gamma}^{AM}(1) = \frac{1}{15}[z_1 z_2 + z_2 z_3 + z_3 z_4 + z_4 z_5 + z_8 z_9 + z_9 z_{10} + \ldots + z_{18} z_{19}].$$

The third autocovariance would be estimated using all observed three day lags, including the three day lags between Friday and Monday:

$$\hat{\gamma}^{AM}(3) = \frac{1}{15}[z_1 z_4 + z_2 z_5 + \underline{z_5 z_8} + z_8 z_{11} + z_9 z_{12} + \underline{z_{12} z_{15}} + z_{15} z_{18} + z_{16} z_{19}].$$

In this method, the denominator for each lag is again the total number of observed data points, as in the the Equal Spacing estimator.

While we focus on these two estimators throughout most of the paper, for comparison, our simulations will also implement two alternatives that are not preferred. The first alternative is the *Parzen estimator*, after Parzen (1963). It is constructed like the Amplitude Modulated estimator, except that we adjust the denominator to equal the number of times we observe

data with the right spacing:

$$\hat{\gamma}^{PZ}(1) = \frac{1}{12}[z_1 z_2 + z_2 z_3 + z_3 z_4 + z_4 z_5 + z_8 z_9 + z_9 z_{10} + ... + z_{18} z_{19}].$$

The third autocovariance would be estimated as:

$$\hat{\gamma}^{PZ}(3) = \frac{1}{8}[z_1 z_4 + z_2 z_5 + z_5 z_8 + z_8 z_{11} + z_9 z_{12} + z_{12} z_{15} + z_{15} z_{18} + z_{16} z_{19}].$$

Finally, we will implement our *Imputation estimator*, which is the Newey-West estimator applied to the filled series, $\{z_t^I\}$, constructed by linearly imputing the missing data. In our example with business daily gasoline prices, the first and third lagged autocorrelations can be estimated as:

$$\hat{\gamma}^{IM}(1) = \frac{1}{21}[z_1 z_2 + ... + z_4 z_5 + z_5 z_6^I + z_6^I z_7^I + z_7^I z_8 + z_8 z_9 + ... + z_{18} z_{19} + z_{19} z_{20}^I + z_{20}^I z_{21}^I]$$

$$\hat{\gamma}^{IM}(3) = \frac{1}{21}[z_1 z_4 + z_2 z_5 + z_3 z_6^I + z_4 z_7^I + z_5 z_8 + z_6^I z_9 + z_7^I z_{10} + z_8 z_{11} + ... + z_{16} z_{19} + z_{17} z_{20}^I + z_{18} z_{21}^I].$$

3 Long-Run Variance of Time Series with Missing Observations

In this section, we formalize our main estimators and describe their asymptotic and finite sample properties.

3.1 Missing Data Structure

Consider a second-order stationary time series $\{z_t\}_{t=1}^{\infty}$ with $\sum_{j=0}^{\infty} |\gamma_z(j)| < \infty$ and $E(z_t) = \mu$ and an indicator series $\{g_t\}_{t=1}^{\infty}$ such that $g_t = 1$ if z_t is observed and $g_t = 0$ if z_t is missing. Throughout the paper, we maintain the assumptions on the missing data structure:

Assumption 1. Independence: *We assume that the underlying series $\{z_t\}$ is independent of the series $\{g_t\}$. In other words, for any positive integer $n < \infty$ and any sequence $t_1, ..., t_n$, the random variable $z \equiv (z_{t_1}, ..., z_{t_n})$ and $g \equiv (g_{t_1}, ..., g_{t_n})$ satisfy the condition that $\Pr(z^{-1}(A) \cap g^{-1}(B)) = \Pr(z^{-1}(A))\Pr(g^{-1}(B))$ for any two n-dimensional Borel sets A, $B \subseteq \mathbb{R}^n$.*

Assumption 2. Existence: $\lim_{T \to \infty}(S/T) = \alpha$ and $\lim_{T \to \infty}[1/S]\sum_{t=j+1}^{T} g_t g_{t-j} = \kappa(j)$ *both exist.*

Assumption 1 requires that the missing process is independent of the underlying data, so that missing data do not induce bias in the parameter estimates. Assumption 2 requires that the fractions of observed converges in probability, and the asymptotic ratio of the number of observed lag j to total number of observations exists. Under these assumptions, we allow very general stochastic or deterministic missing data processes. We give two commonly observed missing data structures as follows:

Example 1. Bernoulli missing: The series $\{g_t\}_{t=1}^{\infty}$ has an i.i.d. Bernoulli distribution, in which each g_t takes value 0 with probability p and value 1 with probability $1 - p$.

Example 2. Cyclically missing: Given a series $\{z_t\}_{t=1}^{\infty}$, we can divide the series into cycles which are each of length $h > 2$. In the first cycle of length h, we have k missing observations for some integer $k < h$. Define the set of time indexes of these missing observations, $S = \{s_1, ..., s_k\}$, where the integers $s_k \in [1, h]$ for all k. For $t \leq h$, $g_t = 0$ if and only if $t \in S$. In a cyclically missing structure, for $t > h$, we have $g_{s_k + hl} = 0$ for all integers $l = 1, 2, ..., \infty$, and $g_t = 1$ otherwise.

The indicator series $\{g_t\}$ is stochastic for Bernoulli missing and deterministic for cyclical missing once the missing pattern is known for any h consecutive elements.

3.2   Newey-West Estimator

First, we review the standard Newey-West estimator that applies to time series without missing observations. Suppose that z_t is continuously observed at $t = 1, ..., T$ with $E(z_t) = \mu$.

We let $\gamma_z(j) = E[(z_t - \mu)(z_{t-j} - \mu)]$ denote the j-th lagged autocovariance. Under the standard assumption that z_t is second-order stationary with $\sum_{j=0}^{\infty} |\gamma_z(j)| < \infty$, we have the standard results that $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (z_t - \mu) \xrightarrow{d} N(0, \Omega)$, where the long-run variance of the underlying process z_t is equal to

$$\Omega = \sum_{j=-\infty}^{\infty} \gamma_z(j). \tag{1}$$

The Newey-West HAC estimator for Ω is given by

$$\hat{\Omega}^{NW} = \hat{\gamma}_z(0) + 2 \sum_{j=1}^{m} w(j, m) \hat{\gamma}_z(j),$$

where $\hat{\gamma}_z(j) = \frac{1}{T} \sum_{t=j+1}^{T} (z_t - \bar{z}_T)(z_{t-j} - \bar{z}_T)$ and $\bar{z}_T = (1/T) \sum_{t=1}^{T} z_t$. In the Newey-West formula, the lagged autocovariances, $\hat{\gamma}_z(j)$, are weighted by the Bartlett kernel, $w(j, m) = 1 - [j/(m+1)]$ for $j \le m$ and $w(j, m) = 0$ otherwise, to ensure a positive semi-definite covariance matrix. Under fairly general technical assumptions, as long as $\lim_{T \to \infty} m(T) = \infty$ and $\lim_{T \to \infty} [m(T)/T^{1/4}] = 0$, we have $\hat{\Omega}^{NW} \xrightarrow{p} \Omega$ (Newey and West, 1987). The choice of optimal bandwidth m is given by Andrews (1991) and Newey and West (1994) who further explore the properties of alternative choices for the bandwidth and kernel and the finite sample properties of these estimators.

3.3 Long-Run Variance of the Underlying Process - Parzen Estimator

In the presence of missing observations, we follow Parzen (1963) and recast the series as an amplitude modulated version of some underlying full series. We define the amplitude modulated series, $\{z^*\}$, as $z_t^* = g_t z_t$. Using the amplitude modulated series $\{z_t^*\}$, Parzen (1963) suggests the following estimator for the autocovariance of the underlying series $\{z_t\}$:

$$\hat{\gamma}_z^{PZ}(j) = \frac{\sum_{t=j+1}^{T} (z_t^* - g_t \bar{z}_T^*)(z_{t-j}^* - g_{t-j} \bar{z}_T^*)}{\sum_{t=j+1}^{T} g_t g_{t-j}},$$

if $\Sigma_{t=j+1}^{T} g_t g_{t-j} > 0$. Dunsmuir and Robinson (1981a) establishes $\hat{\gamma}_z^{PZ}(j) \overset{p}{\to} \gamma_z(j)$ provided that z_t^* is asymptotically stationary.

Under the special case that $\lim_{T\to\infty} \sum_{t=j+1}^{T} g_t g_{t-j} > 0$ for all j, we can use the observed data to construct our *Parzen estimator*, which is a Newey-West type consistent estimator of the long-run variance of the underlying process z_t:

$$\hat{\Omega}^{PZ} = \hat{\gamma}_z^{PZ}(j) + 2\sum_{j=1}^{m} w(j,m)\hat{\gamma}_z^{PZ}(j) \overset{p}{\to} \Omega$$

While this object may be useful in some instances, it is incorrect for inference testing. First, Dunsmuir and Robinson (1981b) study the case in which $w(j,m) = 1$, and point out that $\hat{\Omega}^{PZ}$ may not be positive semi-definite.[2] Secondly, as we further demonstrate, the long-run variance of the underlying process differs from the long-run variance of the observed process. Though the Parzen estimator is formed using observed data only, it is a consistent estimator of the variance of the underlying process. Consequently, inference on the observed data will be invalid if we use the Parzen estimate of the variance.

3.4 Long-Run Variance of the Observed Process

Let $S = \sum_{t=1}^{T} g_t$ be the total number of the observed. The sample mean is given by $\bar{z}_T^* = \frac{1}{S}\sum_{t=1}^{T} z_t^*$. Asymptotic mean and variance of $\bar{z}_T^*$ is given by the following proposition.

Proposition 1. $\bar{z}_T^* \overset{p}{\to} \mu$ *and* $\Omega^* \equiv \lim_{T\to\infty} S \cdot E(\bar{z}_T^* - \mu)^2 = \sum_{j=-\infty}^{\infty} \kappa(j)\gamma_z(j)$.

Proof. Given $E(g_t) = \lim_{T\to\infty}(S/T) = \alpha$ and g_t is independent of z_t, we have $E(z_t^*) = E(g_t)E(z_t) = \alpha\mu$. We can rewrite $\lim_{T\to\infty} \bar{z}_T^* = \lim_{T\to\infty} \frac{1}{S}\sum_{t=1}^{T} z_t^* = \lim_{T\to\infty} \frac{T}{S}\frac{1}{T}\sum_{t=1}^{T} z_t^*$.We know that $\lim_{T\to\infty}(S/T) = \alpha$. By the law of large numbers, we have $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} z_t^* = \alpha\mu$.

---

[2]In our simulations, we implement the estimator using Bartlett kernel weights to maintain comparability with results for ES and AM. However, this does not guarantee that the estimator will be positive semi-definite.

Therefore, $\lim_{T\to\infty} \bar{z}_T^* = \mu$. We also have

$$
\begin{aligned}
S \cdot E(\bar{z}_T^* - \mu)^2 &= [1/S]E\left[\sum_{t=1}^{T} g_t(z_t - \mu)\right]^2 \\
&= [1/S]E\left[\sum_{t=1}^{T}(z_t - \mu)^2 g_t^2 + 2\sum_{j=1}^{T-1}\sum_{t=j+1}^{T}(z_t - \mu)(z_{t-j} - \mu)g_t g_{t-j}\right] \\
&= [T/S]\left[\gamma_z(0)E\left(\frac{1}{T}\sum_{t=1}^{T} g_t^2\right) + 2\sum_{j=1}^{T-1}\gamma_z(j)E\left(\frac{1}{T}\sum_{t=j+1}^{T} g_t g_{t-j}\right)\right].
\end{aligned}
$$

Define $\kappa(j) \equiv \lim_{T\to\infty}[1/S]\sum_{t=j+1}^{T} E(g_t g_{t-j})$, the share of times lag j is observed. Given $\lim_{T\to\infty}(T/S) = 1/\alpha$ and $E\left[(1/T)\sum_{t=j+1}^{T} g_t g_{t-j}\right] = \lim_{T\to\infty}(1/T)\sum_{t=j+1}^{T} g_t g_{t-j} = \alpha\kappa(j)$, we have

$$
\Omega^* = \lim_{T\to\infty} S \cdot E(\bar{z}_T^* - \mu)^2 = \sum_{j=-\infty}^{\infty} \kappa(j)\gamma_z(j). \tag{2}
$$

$\square$

Therefore, the long-run variance of the observed amplitude modulated process, i.e., Ω^*, is a weighted sum of the original autocovariances, with the weights being the asymptotic ratio of the number of the observed lags to the total number of the observed, S. Comparing equations (1) and (2), when z_t is observed at all t, i.e., $g_t = 1$ for all t, then $\Omega^* = \Omega$. In the presence of missing observations, if all autocovariances are positive, we have $\kappa(j) \leq 1$. Then the long-run variance of the amplitude modulated process is always weakly smaller than the long-run variance of the underlying process, $\Omega^* \leq \Omega$.

3.4.1 Amplitude Modulated Estimator

To estimate Ω^* in finite samples with S observed, a natural candidate is given by the following proposition.

Proposition 2. *A consistent estimator of Ω^* is given by*

$$\hat{\Omega}^* = \hat{\gamma}_{z^*}(0) + 2\sum_{j=1}^{T-1}\hat{\gamma}_{z^*}(j)$$

where $\hat{\gamma}_{z^*}(j) = [1/S]\sum_{t=j+1}^{T}(z_t^* - g_t\bar{z}_T^*)(z_{t-j}^* - g_{t-j}\bar{z}_T^*)$.

Proof. We note that

$$
\begin{aligned}
\hat{\gamma}_{z^*}(j) &= [1/S]\sum_{t=j+1}^{T}(z_t^* - g_t\bar{z}_T^*)(z_{t-j}^* - g_{t-j}\bar{z}_T^*) \\
&= \frac{T}{S}\frac{1}{T}\sum_{t=j+1}^{T}(z_t - \bar{z}_T^*)(z_{t-j} - \bar{z}_T^*)g_t g_{t-j}
\end{aligned}
$$

Since $\lim_{T\to\infty} T/S = 1/\alpha$, $\lim_{T\to\infty} E(z_t - \bar{z}_T^*)(z_{t-j} - \bar{z}_T^*) = \gamma_z(j)$ and $\lim_{T\to\infty} E(g_t g_{t-j}) = \alpha\kappa(j)$ we have

$$\hat{\gamma}_{z^*}(j) \xrightarrow{p} \kappa(j)\gamma(j).$$

Therefore, $\hat{\Omega}^* \xrightarrow{p} \Omega^*$. $\qquad\square$

However, $\hat{\Omega}^*$ is not guaranteed to be positive semi-definite, which is not desirable for inference. We can use a kernel-based method to ensure the covariance estimator is positive semi-definite:

$$\hat{\Omega}^{AM} = \hat{\gamma}_{z^*}(0) + 2\sum_{j=1}^{m}w(j,m)\hat{\gamma}_{z^*}(j).$$

We follow Newey and West (1987) and illustrate with the most commonly used kernel, the Bartlett kernel.

Proposition 3. *Using the Bartlett kernel,* $w(j,m) = 1 - [j/(m+1)]$ *if* $j \leq m$ *and* $w(j,m) = 0$ *if* $j > m$, *suppose (i) the bandwidth* m *satisfies* $\lim_{T\to\infty} m(T) = +\infty$ *and* $\lim_{T\to\infty}[m(T)/T^{1/4}] = 0$. *Then* $\hat{\Omega}^{AM}$ *is positive semi-definite and* $\hat{\Omega}^{AM} \xrightarrow{p} \Omega^*$.

Proof. We follow proofs of Theorems 1 and 2 in Newey and West (1987) by defining $h_t \equiv z_t^* - g_t\mu$, $\hat{h}_t \equiv z_t^* - g_t\bar{z}_T^*$ and replace all T in the denominator in Newey and West (1987) with S. $\square$

Our estimator $\hat{\Omega}^{AM}$ is almost equivalent to applying the Newey-West estimator to the amplitude modulated series. However, we make two minor modifications to the components $\hat{\gamma}_{z^*}(j) = [1/S]\sum_{t=j+1}^{T}(z_t^* - g_t\bar{z}_T^*)(z_{t-j}^* - g_{t-j}\bar{z}_T^*)$. First, we subtract z_t^* by $g_t\bar{z}_T^*$ instead of $E(g_t)\bar{z}_T^*$, so that the difference $(z_t^* - g_t\bar{z}_T^*)$ equals zero for unobserved data. In the case of a mean-zero series, this modification would not be required. Secondly, since we want to use $\hat{\Omega}^{AM}$ to make inferences about the mean of the observed process $\bar{z}_T^*$, we divide the sum by S instead of T so that our inference procedure remains consistent.

3.4.2 Equal Spacing Estimator

Instead of casting the time series with missing observations as an amplitude modulated process, an alternative method is to ignore the missing observations and treat the data as equally spaced over time. We define the function $\iota(t)$ as the mapping from time index t to the new equal spacing time domain s: $\iota(t) = \sum_{\ell=1}^{t} g_\ell$. We use this mapping to relabel the time indices of the observed values from the series $\{z_t\}$ to create the series $\{z_s^{ES}\}$ for $s = 1, ..., \iota(T)$, in the equal spacing time domain. The sample mean of the observed series is given by $\bar{z}_T^{ES} = \frac{1}{S}\sum_{s=1}^{S} z_s^{ES}$.

Proposition 4. $\bar{z}_T^{ES} \overset{p}{\to} \mu$ *and* $\Omega^{ES} \equiv \lim_{T\to\infty} S \cdot E(\bar{z}_T^{ES} - \mu)^2 = \Omega^*$.

Proof. We let $\Delta_s^{j^{ES}} \equiv \iota^{-1}(s) - \iota^{-1}(s - j^{ES})$ be a function that maps the time gap (j^{ES}) between z_s and $z_{s-j^{ES}}$ in the equal spacing domain to the time gap (j) in the time domain of the underlying process $\{z_t\}$. Using the indicator function $\mathbf{I}(\cdot)$, we define $\lambda^{j^{ES}}(j) = \lim_{T\to\infty} \frac{1}{S}\sum_{s=1}^{S} \mathbf{I}(\Delta_s^{j^{ES}} = j)$, which equals the frequency that the observed lag j^{ES} maps to lag j in the original time domain. Then we can rewrite the Equal Spacing autocovariance in

terms of the autocovariance of the underlying process:

$$
\gamma_{z^{ES}}(j^{ES}) = \lim_{T \to \infty} E(z_s - \bar{z}_T^{ES})(z_{s-j^{ES}} - \bar{z}_T^{ES})
$$

$$
= \sum_{j=-\infty}^{\infty} \lambda^{j^{ES}}(j)\gamma_z(j)
$$

Applying the same standard results as in Equation 1 to the equal spacing series, we have:

$$
\Omega^{ES} = \sum_{j^{ES}=-\infty}^{\infty} \gamma_{z^{ES}}(j^{ES})
$$

$$
= \sum_{j^{ES}=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \lambda^{j^{ES}}(j)\gamma_z(j)
$$

$$
= \sum_{j=-\infty}^{\infty} \left(\sum_{j^{ES}=-\infty}^{\infty} \lambda^{j^{ES}}(j) \right) \gamma_z(j)
$$

$$
= \sum_{j=-\infty}^{\infty} \kappa(j)\gamma_z(j) = \Omega^*.
$$

The second to last equation holds because

$$
\sum_{j^{ES}=-\infty}^{\infty} \lim_{T \to \infty} \frac{1}{S} \sum_{s=1}^{S} \mathbf{I}(\Delta_s^{j^{ES}} = j) = \lim_{T \to \infty} \frac{1}{S} \sum_{t=1}^{T} g_t g_{t-j}.
$$

$\square$

To estimate Ω^* using the equally spaced series in finite samples, we can use

$$
\hat{\Omega}^{ES} = \hat{\gamma}_{z^{ES}}(0) + 2 \sum_{j=1}^{m} w(j^{ES}, m)\hat{\gamma}_{z^{ES}}(j^{ES})
$$

where

$$
\hat{\gamma}_{z^{ES}}(j^{ES}) = \frac{1}{S} \sum_{j=1}^{S-1} (z_s^{ES} - \bar{z}_T^{ES})(z_{s-j}^{ES} - \bar{z}_T^{ES}).
$$

Proposition 5. $\hat{\Omega}^{ES}$ is PSD and $\hat{\Omega}^{ES} \xrightarrow{p} \Omega^*$.

Proof. Positive semi-definiteness of $\hat{\Omega}^{ES}$ can be established using same argument in the proof of Theorem 1 in Newey and West (1987). Using their notation, we let $\hat{h}_s = z_s^{ES} - \bar{z}_T^{ES}$. To prove consistency, since

$$\hat{\gamma}_{z^{ES}}(0) + 2 \sum_{j^{ES}=1}^{s(T)-1} \hat{\gamma}_{z^{ES}}(j^{ES}) = \hat{\gamma}_{z^*}(0) + 2 \sum_{j=1}^{T-1} \hat{\gamma}_{z^*}(j)$$

and $w(j, m) \xrightarrow{p} 1$ for all j, $\hat{\Omega}^{ES}$ and $\hat{\Omega}^{AM}$ estimators are asymptotically equivalent. We know $\hat{\Omega}^{AM} \xrightarrow{p} \Omega^*$ by proposition 3, and hence, $\hat{\Omega}^{ES} \xrightarrow{p} \Omega^*$.

The Equal Spacing estimator is particularly simple to implement, because it only requires relabeling the time index of a series with missing observations to ignore the gaps and treat the data as equally spaced over time. Once this is done, the Equal Spacing estimator amounts to applying the standard Newey-West estimator to the equally spaced series. $\qquad\square$

3.5   Finite Samples

Although $\hat{\Omega}^{AM}$ and $\hat{\Omega}^{ES}$ are both asymptotically consistent, finite sample performance might differ due to different weighting on autocovariance estimators. We use the standard mean squared error (MSE) criterion to to evaluate the performance of the estimator $\hat{\Omega}^i$ where $i \in \{AM, ES\}$.

$$
\begin{aligned}
MSE(\hat{\Omega}^i) &= Bias^2(\hat{\Omega}^i) + Var(\hat{\Omega}^i) \\
&= \left[E(\hat{\Omega}^i - \Omega^*) \right]^2 + E[(\hat{\Omega}^i - \bar{\Omega}^i)^2]
\end{aligned}
$$

where $\bar{\Omega}^i = E(\hat{\Omega}^i)$. Consider the case that $m^{AM} = m^{ES} \equiv m$. The lag length in the original time domain is weakly greater than that in the equal spacing domain: $j = \iota^{-1}(s) - \iota^{-1}(s - j^{ES}) \geq j^{ES}$. Under the same fixed bandwidth m, there are two main differences between $\hat{\Omega}^{AM}$ and $\hat{\Omega}^{ES}$. First, since the kernel weight is decreasing in the lag length for the Newey-West estimator, $\hat{\Omega}^{ES}$ assigns weakly higher weight on all autocovariance estimators

compared to $\hat{\Omega}^{AM}$. To see this more explicitly,

$$\hat{\Omega}^{ES} = \sum_{j^{ES}=1}^{m} \frac{w(j^{ES}, m)}{S} \sum_{s=j^{ES}+1}^{S} (z_s - \bar{z}_T^*)(z_{s-j^{ES}} - \bar{z}_T^*)$$

To write it in the original time domain, we have

$$\hat{\Omega}^{ES} = \sum_{j^{ES}=1}^{m} \frac{w(j^{ES}, m)}{S} \sum_{t=j+1}^{T} (z_t - \bar{z}_T^*)(z_{t-j} - \bar{z}_T^*)$$

where $t = \iota^{-1}(s)$ and $j = \iota^{-1}(s) - \iota^{-1}(s - j^{ES}) \geq j^{ES}$. We compare $\hat{\Omega}^{ES}$ with $\hat{\Omega}^{AM}$,

$$\hat{\Omega}^{AM} = \sum_{j=1}^{m} \frac{w(j, m)}{S} \sum_{t=j+1}^{T} (z_t - \bar{z}_T^*)(z_{t-j} - \bar{z}_T^*) g_t g_{t-j}.$$

When $g_t g_{t-j} = 1$, we have $w(j^{ES}, m) \geq w(j, m)$ since the weighting function decreases in lag length and $j \geq j^{ES}$. Therefore, given the same bandwidth, $\hat{\Omega}^{ES}$ puts weakly more weight than $\hat{\Omega}^{AM}$ on each observed pairwise product $(z_t - \bar{z}_T^*)(z_{t-j} - \bar{z}_T^*)$. Second, for the same fixed bandwidth m, $\hat{\Omega}^{AM}$ only estimates autocovariance with lag length up to m in the original time domain, while $\hat{\Omega}^{ES}$ also includes observations for lags greater than m the original time domain. These two differences have different implications on the relative variance and bias of the two estimators.

As discussed in den Haan and Levin (1997), Newey-West type kernel-based estimators suffer from three sources of finite sample bias. First, the summation in the autocovariance estimator is divided by the sample size, instead of the actual number of observed lags. We expect this source of bias to be more severe for $\hat{\Omega}^{ES}$ because $\hat{\Omega}^{ES}$ includes higher-order lags that are not included in $\hat{\Omega}^{AM}$ and puts more weight on these high-order biased lagged autocovariance estimators. However, this bias decreases rapidly as the sample size increases. Second, the kernel-based method assigns zero weights to lags with orders greater than T. This source of bias is the same for $\hat{\Omega}^{ES}$ and $\hat{\Omega}^{AM}$.

The third and most significant source of bias is driven by the fact that kernel-based estimators under-weight the autocovariance estimators. They assign weights to autocovariance estimators that are less than unity and are declining toward zero with increasing lag order j. Compared with the long-run variance of the amplitude modulated series, Ω^*, the bias of $\hat{\Omega}^{AM}$ arising from this source is given by

$$Bias(\hat{\Omega}^{AM}) = \sum_{j=-(T-1)}^{T-1} [1 - w(j,m)] \gamma_{z^*}(j).$$

For a fixed bandwidth, the higher the serial correlation, the more severe the bias. The estimator $\hat{\Omega}^{ES}$ can reduce this kernel-based bias because $\hat{\Omega}^{ES}$ always assigns weakly higher (or closer to unitary) weight to all autocovariance estimators as compared to $\hat{\Omega}^{AM}$.

For variance of the estimators, we always have $Var(\hat{\Omega}^{ES}) > Var(\hat{\Omega}^{AM})$ because $\hat{\Omega}^{ES}$ includes more high-order lags that are relatively poorly estimated. Therefore, the tradeoff between variance and bias determines the relative finite sample performance of $\hat{\Omega}^{AM}$ and $\hat{\Omega}^{ES}$. The previous discussion uses a fixed bandwidth. Andrews (1991) and Newey and West (1994) propose data-dependent choice of the bandwidth that aims to optimize the mean-variance tradeoff. We apply the automatic bandwidth selection procedure proposed by Newey and West (1994) to the AM and ES processes. As we will demonstrate using Monte-Carlo simulations, under both fixed and automatic bandwidth selection, for small sample size and low autocorrelation, $MSE(\hat{\Omega}^{ES}) > MSE(\hat{\Omega}^{AM})$. For moderate sample size or high autocorrelation, we always have $MSE(\hat{\Omega}^{AM}) > MSE(\hat{\Omega}^{ES})$.

4   Regression Model with Missing Observations

We can apply asymptotic theory developed in the previous section to a regression model with missing observations. Suppose we have the time series regression, where y_t and u_t are scalars, x_t is a $k \times 1$ vector of regressors, and β is a $k \times 1$ vector of unknown parameters. Suppose

further that $\left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1} \xrightarrow{p} \Sigma_{xx}^{-1}$ and $\mathbb{E}(u_t|x_t) = 0$, but the u_t's have conditional het-eroskedasticity and are possibly serially correlated. In the presence of missing observations, we let $g_t = 1$ if y_t and all components of x_t are observed and $g_t = 0$ if y_t or any component of x_t is missing. Then we can re-express the regression in terms of amplitude modulated processes,

$$y_t^* = x_t^* \beta + u_t^*, \ t = 1, \ldots, T,$$

where $y_t^* = g_t y_t$, $x_t^* = g_t x_t$ and $u_t^* = g_t u_t$. We require the orthogonality condition, $\mathbb{E}(u_t^*|x_t^*) = 0$. The standard result for the OLS estimator is given by

$$\hat{\beta}^{AM} - \beta = \left(\sum_{t=1}^{T} x_t^* x_t^{*'}\right)^{-1} \left(\sum_{t=1}^{T} x_t^* u_t^*\right). \tag{3}$$

Alternatively, without recasting the series as an amplitude modulated process, we ignore all observations for which $g_t = 0$ and assume all observed values are equally spaced in time. Therefore, the estimated regression becomes

$$y_s^{ES} = x_s^{ES} \beta + u_s^{ES}, \ s = 1, \ldots, S$$

and

$$\hat{\beta}^{ES} - \beta = \left(\sum_{s=1}^{S} x_s^{ES} x_s^{ES'}\right)^{-1} \left(\sum_{s=1}^{S} x_s^{ES} u_s^{ES}\right). \tag{4}$$

Comparing equations 3 and 4, we can easily see that AM and ES give the same coefficient estimates:

$$\hat{\beta}^{AM} = \hat{\beta}^{ES} \equiv \hat{\beta}.$$

We normalize $\hat{\beta}$ using the number of observed data, S, and then we have

$$\sqrt{S}(\hat{\beta} - \beta) = \left(\frac{1}{S}\sum_{t=1}^{T} x_t^* x_t^{*'}\right)^{-1} \left(\frac{1}{\sqrt{S}}\sum_{t=1}^{T} x_t^* u_t^*\right).$$

21

Given that $\left(\frac{1}{T}\sum_{t=1}^{T}x_t x_t'\right)^{-1} \xrightarrow{p} \Sigma_{xx}^{-1}$ in the absence of missing observations and x_t and g_t are independent, $\left(\frac{1}{S}\sum_{t=1}^{T}x_t^* x_t^{*'}\right)^{-1}$ also converges in probability. We let $\left(\frac{1}{S}\sum_{t=1}^{T}x_t^* x_t^{*'}\right)^{-1} \xrightarrow{p} \Sigma_{x^*x^*}^{-1}$. Using the notation from the previous section, we define $z_t \equiv x_t u_t$ and let $z_t^* \equiv g_t z_t$ denote the amplitude modulated series and z_s^{ES} denote the ES series. Then we have

$$\bar{z}_T^* \equiv \frac{1}{S}\sum_{t=1}^{T}z_t^* = \frac{1}{S}\sum_{t=1}^{S}z_s^{ES}.$$

We know $E(z_t) = E(\bar{z}_T^*) = 0$ using the orthogonality condition.

Proposition 6. *The asymptotic distribution of the OLS estimator is given by*

$$\sqrt{S}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma_{x^*x^*}^{-1}\Omega^*\Sigma_{x^*x^*}^{-1}),$$

where $\Omega^* = \sum_{j=-\infty}^{\infty}\kappa(j)\gamma_z(j)$ *and* $\kappa(j) = \lim_{S\to\infty}\frac{1}{S}\sum_{t=j+1}^{T}E(g_t g_{t-j})$.

To estimate $\Sigma_{x^*x^*}^{-1}$, we can use

$$\hat{\Sigma}_{x^*x^*}^{-1} = \left(\frac{1}{S}\sum_{t=1}^{T}x_t^* x_t^{*'}\right)^{-1} = \left(\frac{1}{S}\sum_{s=1}^{S}x_s^{ES}x_s^{ES'}\right)^{-1} \xrightarrow{p} \Sigma_{x^*x^*}^{-1}.$$

Proposition 7. *We define*

$$\hat{\Omega}^{AM} = \hat{\Gamma}_0^{AM} + \sum_{j=1}^{m}w(j,m)[\hat{\Gamma}_j^{AM} + \hat{\Gamma}_j^{AM'}],$$

where $\hat{\Gamma}_j^{AM} = (1/S)\sum_{t=1}^{T}z_s^* z_{s-j}^{*'}$,

$$\hat{\Omega}^{ES} = \hat{\Gamma}_0^{ES} + \sum_{j=1}^{m}w(j,m)[\hat{\Gamma}_j^{ES} + \hat{\Gamma}_j^{ES'}],$$

where $\hat{\Gamma}_j^{ES} = (1/S)\sum_{s=1}^{S}z_s^{ES}z_{s-j}^{ES'}$. *Then we have* $\hat{\Omega}^{AM} \xrightarrow{p} \Omega^*$ *and* $\hat{\Omega}^{ES} \xrightarrow{p} \Omega^*$. *For inferences, the t-statistic based on* $\hat{\Omega}^{AM}$ *is given by*

$$t_k^{AM} = \frac{\hat{\beta}_k - \beta_0}{\sqrt{\hat{V}_{kk}^{AM}/(S-k)}} \xrightarrow{d} N(0,1),$$

where $\hat{V}_{kk}^{AM}$ is the (k,k)-th element of $\hat{V}^{AM} = \hat{\Sigma}_{x^*x^*}^{-1}\hat{\Omega}^{AM}\hat{\Sigma}_{x^*x^*}^{-1}$. Alternatively, the t-statistic based on $\hat{\Omega}^{ES}$ is given by

$$t_k^{ES} = \frac{\hat{\beta}_k - \beta_0}{\sqrt{\hat{V}_{kk}^{ES}/(S-k)}} \xrightarrow{d} N(0,1),$$

where $\hat{V}_{kk}^{ES}$ is the (k,k)-th element of $\hat{V}^{ES} = \hat{\Sigma}_{x^*x^*}^{-1}\hat{\Omega}^{ES}\hat{\Sigma}_{x^*x^*}^{-1}$.

5   Simulation

In the Monte Carlo simulations that follow, we study the properties of our Amplitude Modulated and Equal Spacing estimators using a simple location model. To evaluate these estimators under a variety of circumstances, we generate data with various levels of autocorrelation and for a range of sample sizes. We test our estimators under the two missing structures described in Section 3.1, the Bernoulli missing structure and the deterministic cyclically missing structure. We implement the estimators using a standard fixed bandwidth for our benchmark results, and also provide results that implement the automatic bandwidth selection procedure proposed by Newey and West (1994). Our primary evaluation criteria are the empirical rejection probability of the test, and the power of the test against an appropriate alternative.

5.1 Data Structure

The inference procedures are tested on a simulated data series $\{y_t\}$ that is generated using a simple location model:

$$y_t = \beta + \epsilon_t$$

$$\epsilon_t = \phi\epsilon_{t-1} + \eta_t$$

$$\text{where } \eta_t \text{ i.i.d. } \sim \mathcal{N}(0,1) \text{ and } \epsilon_0 = 0$$

For each of $N = 100,000$ iterations, we use $\beta = 0$ and generate a data series $\{y_1, ..., y_{T^{max}}\}$ with a sample size of $T^{max} = 24,000$. Since we run tests over a range of sample sizes for each estimator, we use the first T observations in each iteration for $T \in \{120, 360, 1200, 4800, 12000, 24000\}$. To test these methods for a range of autocorrelation parameters, ϕ, we generate data separately for $\phi \in \{0, 0.3, 0.5, 0.7, 0.9\}$. The regressor in this model is a constant 1, and we conduct simulation exercises under different missing structures for the dependent variable $\{y_t\}$, which are described below. For each iteration, we generate a series $\{g_t\}$, which indicates for each t whether y_t is observed. Finally, we generate the series $\{y_t^*\}$, where $y_t^* = g_t y_t$.

Given this data, we estimate the parameter of interest, β^i, and the estimator for the covariance matrix, $\hat{\Omega}^i$, for each estimator $i \in \{NW, ES, AM\}$. We also perform simulations for two additional methods. First, we implement the imputation method, in which the missing y_t are linearly imputed before the standard Newey-West estimator is applied to the filled series $\{y_t^I\}$. Secondly, we implement the Parzen estimator from Section 3.3. Since the estimator $\hat{\Omega}^{PZ}$ is not positive semi-definite, we calculate the rejection rate using the number of rejections divided by the number of simulations in which $\hat{\Omega}^{PZ} > 0$.

We use these estimators to calculate the t-statistic, t_β^i, used for a test of the null hypothesis $H_0 : \beta = 0$ against the two-sided alternative, $H_a : \beta \neq 0$. We choose a 5% level of significance and reject the null hypothesis when $|t_\beta^i| > 1.96$. For the standard estimations,

24

we use a fixed bandwidth of $m = 4(T/100)^{(2/9)}$. We also apply the automatic bandwidth selection procedure of Newey and West (1994) to the Newey-West, Equal Spacing, Amplitude Modulating, and Imputation methods.

Results are reported for simulation exercises under a variety of sampling schemes. Our benchmark sampling scheme is one with a Bernoulli missing structure as described in Example 1 of Section 3.1. For these simulations, the series $\{g_t\}$ has an i.i.d. Bernoulli distribution with fixed probability of missing, $p = 6/12$. For comparison, we also provide two variants in which the probability of missing is set to $4/12$ and $8/12$.

We also simulate data under four data structures with cyclically missing observations, as described in Example 2 of Section 3.1. For these, we choose a cycle length of 12, with 6 or 8 observations missing each cycle. In these simulations, the missing structure is cyclical in that we generate a single pattern of missing observations for the first cycle, and apply this same pattern to every cycle of 12 observations. Additionally, the pattern is deterministic in the sense that we apply the same pattern of missing observations for all iterations in the simulation. This sampling structure reflects the potential application of these methods to monthly data with missing observations. For example, because commodities futures contracts expire only some months of the year, the data on monthly commodities returns will have the same missing pattern each year. Another common application is for daily financial data, in which the same 2 weekend days are missing in each cycle of 7 days.

Under a deterministic cyclical missing structure, it is possible to have cases for which certain lagged autocovariances in the original time series domain are never observed. As we noted in the introduction, the popular statistical software Stata forbids implementation of the Amplitude Modulated estimator under this case, even when using the "force" command. Yet, our theoretical results do not require all the lags to be observed to generate asymptotically correct inference using the ES and AM estimators. Consequently, we perform simulations for deterministic cyclical missing structures under both cases: all lags are observed at least once, or some lags are never observed. We show that the finite sample performance does not

differ much between these two cases, and neither case differs much from the results under the Bernoulli missing structure.

In our first two cyclical missing structures, we set the cyclical pattern of observed data such that each lagged autocovariance can be estimated from the observed data. In our cyclical structure with 6 of 12 missing, we observe $\{z_3, z_6, z_8, z_9, z_{10}, z_{11}\}$, and then observe the same pattern of observations for each cycle of length 12. In our next cyclical structure, we have 8 of 12 missing. We observe $\{z_3, z_4, z_7, z_9\}$ in the first cycle, and the same pattern for each cycle after that. For our final two cyclical missing structures, we require the observed data to be spaced within the cycle such that at least one lag is never observed. In our structure with 6 of 12 missing, we observe $\{z_1, z_3, z_5, z_8, z_{10}, z_{12}\}$ in the first cycle. Under this structure, the sixth lag is never observed. For our cyclical structure with 8 of 12 missing, we observe $\{z_2, z_3, z_6, z_{12}\}$, so that the fifth lag is never observed.

5.2 Evaluation Criteria

The primary evaluation criteria for these estimators is the empirical rejection probability of the tests. The empirical rejection probability measures the likelihood that null hypothesis is rejected when it is in fact true (Type I error). Each iteration of the Monte Carlo simulation represents one hypothesis test, and the reported rejection probability reflects the fraction of iterations for which the t-statistic was large enough in magnitude to reject the null hypothesis.

We also provide measures of the power of the test, as well as measures of the bias and variance of the estimators. The power of the test measures the probability that the null hypothesis is rejected when the alternative hypothesis is true. Since we find empirical rejection probabilities that can be much higher than the 0.05 benchmark, we calculate the size-adjusted power for ease of comparability. Following Ibragimov and Müller (2010), we set the alternative hypothesis to $H_a : \beta_a = 4/\sqrt{T(1 - \phi^2)}$. To calculate the power, we first calculate $t^i_{\beta_a}$, which is analogous to t^i_β for each i, except that we subtract β_a instead of β_0 in

the numerator. For example,

$$t_{\beta_a}^{NW} = \frac{\hat{\beta}^{NW} - \beta_a}{\sqrt{\hat{V}_{kk}^{NW}/T}}.$$

Next, we calculate an adjusted critical value, $t_{0.05}^{crit}$, which is the t-statistic at the 5th percentile of our simulations. This value is equal to the critical value for which the empirical rejection probability would have been exactly 0.05 under our simulation procedure. To calculate the size-adjusted power, we calculate $t_{\beta_a}^i$ under the alternative hypothesis above and reject when $|t_{\beta_a}^i| > t_{0.05}^{crit}$.

Finally, in order to understand the finite sample performance, we study the empirical mean squared error of our estimators by decomposing it into the empirical variance and bias. Under the benchmark Bernoulli case, we first calculate the value of Ω^* under our data generating process as:

$$\begin{aligned}
\Omega^* &= \lim_{T \to \infty} (T/S) \sum_{j=-\infty}^{\infty} \gamma_j E(g_t g_{t-j}) \\
&= (1/p)\left(p + 2\sum_{j=1}^{\infty} p^2 \phi^j\right) Var(\epsilon_t) \\
&= \left(1 + \frac{2p\phi}{1-\phi}\right)\left(\frac{1}{1-\phi^2}\right)
\end{aligned}$$

where p is the probability of being observed. The second equation follows because (1) $\lim_{T\to\infty} T/S = 1/p$; (2) $E(g_t g_{t-j}) = p$ if $j = 0$ and $E(g_t g_{t-j}) = p^2$ if $j \geq 1$. The third equation holds because $Var(\epsilon_t) = 1/(1 - \phi^2)$. Returning to the MSE decomposition, we have:

$$\begin{aligned}
\widehat{MSE} &= \widehat{Bias}^2 + \widehat{Variance} \\
&= \left(\hat{\bar{\Omega}}^i - \Omega^*\right)^2 + \frac{1}{N}\sum_{c=1}^{N}\left(\hat{\Omega}_c^i - \hat{\bar{\Omega}}^i\right)^2,
\end{aligned}$$

where $\hat{\bar{\Omega}}^i = (1/N) \sum_{c=1}^{N} \hat{\Omega}_c^i$ is the sample mean of all the covariance estimators and c indexes the $N = 100,000$ iterations. Note that for $i = NW$, we have $p = 1$. We use these measures to study the finite sample properties of our estimators, especially to compare the AM and ES estimators.

The primary findings are reported in Tables 2 through 7.[3]

5.3 Results

Table 2 provides the rejection probabilities for our two main estimators, the Equal Spacing estimator and the Amplitude Modulated estimator. We provide results under a fixed bandwidth and automatic bandwidth selection for each estimator, and for comparison purposes, present results for the application of the Newey-West estimator applied to the full series without missing observations. Our benchmark missing data structure is the Bernoulli missing structure in which observations are missing with probability $p = 1/2$. We focus on the results for the simulation with $T = 360$, and also provide results for a large and very large sample size ($T = 1200$ or 24000).

Our simulation results provide evidence that the ES and AM estimators perform well in finite samples. As is well known in the HAC literature, we find that the empirical rejection probability can be a bit higher than 5.0 for small samples, even when there is no autocorrelation. In addition, when the autocorrelation parameter is high, there can be quite a bit of overrejection even for very large sample sizes ($T = 24000$). However, we do find that the rejection probability is falling towards 5.0 as the sample size increases.

We also find evidence that our ES and AM estimators are well-behaved under deterministic cyclically missing structures. In Table 3, we see little difference in the rejection rates under each of our three data structures: randomly missing under a Bernoulli structure, deterministic cyclical missing when all lags are observed, and deterministic cyclical missing when some lags are unobserved.

---

[3]Full simulation results are available upon request.

Table 2 also provides the empirical rejection probabilities for the Parzen and Imputation estimators for $T = 360$ and $T = 24000$. As expected, the higher serial correlation induced by the imputation procedure results in extremely high rejection rates as compared to the ES and AM estimators. We can also see in Table 2 that the results for the Parzen estimator substantiate our argument that this estimator cannot be used for robust inference for series with missing observations. In our simulation with $\phi = 0$ and $T = 360$, we found 20 instances (out of 100,000) in which the non-PSD Parzen estimator returned a negative estimate of the variance. Additionally, we find that the rejection probability is generally decreasing in the sample size but is U-shaped with respect to the autocorrelation, and often underrejects for low levels of autocorrelation.[4]

Next we turn to a comparison of the finite sample properties of the ES and AM estimators. In the results for the fixed bandwidth estimators in Table 4, we find that for $T = 360$, the AM estimator is preferred for autocorrelation parameters $\phi \leq 0.3$, while the ES estimator is preferred for series with higher autocorrelation. For larger samples, the ES estimator is preferred for a larger range of ϕ, so that for $T = 24000$, we have that the ES estimator is preferred for all the simulations with nonzero autocorrelations.

To better understand these results, we turn to Table 5, which reports the empirical bias and variance for these two estimators under our benchmark simulations. As expected, when using the same fixed bandwidth, the ES estimator has a higher variance than the AM estimator for each sample size and autocorrelation. As discussed in Section 3.5, this is because compared to $\hat{\Omega}^{AM}$, the $\hat{\Omega}^{ES}$ estimator includes more high-order lags that are relatively poorly estimated.

With regard to the bias, the ES estimator has better performance for higher autocorrelation parameters, though this effect is mitigated and sometimes reversed in small samples. The poor small sample performance is driven by the first source of bias discussed in Section

---

[4]Interestingly, the test using the PZ estimator is well-behaved when the autocorrelation is 0. This is consistent with our theoretical results, because when there is no autocorrelation, we have that the long-run variance of the underlying and observed series are asymptotically equivalent. Consequently, we have that when there is no autocorrelation, $\hat{\Omega}^{PZ}$ and $\hat{\Omega}^{AM}$ are asymptotically equivalent as well.

3.5, that the summation in the autocovariance estimator is divided by the sample size rather than the actual number of observed lags. This bias declines rapidly as the sample size increases. In contrast, the bias behind the overrejection for high autocorrelations is driven by the underweighting of high-order lagged autocovariances. Since the ES estimator places a higher weight on high-order autocovariances, it has lower bias than the AM estimator when the autocorrelation is high. As the sample size grows and the first source of bias becomes less important, the ES estimator is preferred for a larger range of autocovariance parameters.

This bias and variance tradeoff changes when we implement automatic bandwidth selection. The results in Table 4 indicate that under this procedure, the AM estimator has a lower rejection probability and is thus preferred at all but the highest level of autocorrelation, for every sample size. To provide further context for this result, Table 6 reports the average selected bandwidth for each simulation. We know from den Haan and Levin (1997) that using a higher bandwidth will increase the variance of the estimator while decreasing the bias. Given that the AM estimator has a lower variance than the ES estimator when using a fixed bandwidth, it is not surprising that the automatic bandwidth selection typically chooses a higher bandwidth for the AM estimator than for the ES estimator. The incremental improvement in the bias between the fixed and automatic bandwidth selection is larger for the AM estimator than for the ES estimator. Consequently, under automatic bandwidth selection, the ES estimator is only preferred for extremely high autocorrelation, when the bias of the AM estimator is much higher than that of the ES estimator.

Turning to the size-adjusted power, we find in table Table 4 that the power of the two estimators is roughly equivalent. Comparing our two main estimators, we have that the power of the AM estimator is generally stronger than that of the ES estimator. Just as for the Newey-West results, we have that the power falls as the autocorrelation increases, and that the size-adjusted power does not vary as the sample size increases. Unsurprisingly, we can see that the power of the test under the missing structure is weaker than for the full series, due to the smaller observed sample size. This effect is mitigated at high autocorrelation,

however, and we can see that under very high autocorrelation, the power is roughly equal for the Newey-West application to the full series and for the application of our two estimators to the series with missing observations.

Finally, Table 7 presents results for varying fractions of missing observations. At low autocorrelation, the rejection rate increases as the fraction of missing observations increases. This is likely driven by the first source of finite sample bias discussed in Section 3.5, which gets worse as the fraction of missing observations increases. In contrast, at high autocorrelation, the rejection rate falls as the fraction of missing observations increases. This effect is likely due to the fact that when a higher fraction of observations is missing, the observed process is less persistent, and the estimators are better able to overcome the underweighting of the higher order autocovariances. Putting these two effects together, we have that the AM estimator is preferred for a larger range of autocorrelation parameters when a higher fraction of data is missing. This is consistent with our previous finding, that the AM estimator is preferred when the serial correlation is low.

Overall, these simulation results are consistent with our theoretical findings. We show that the ES and AM estimators both are well-behaved for random and deterministic missing structures. In general, the ES and AM estimators are preferred to imputation methods, which may induce serial correlation in the imputed series, resulting in more severe bias and overrejection. In finite samples, we find that for the same fixed bandwidth, the ES estimator is generally less biased than the AM estimator, but has larger variance. Consequently, the ES estimator is preferred when autocorrelation is high, as the bias will dominate the mean squared error in these cases. Conversely, when autocorrelation is low, variance dominates the mean squared error, and the AM estimator is preferred.

6 Empirical Application: Recursive Tests for a Positive Sample Mean

In this section, we present an application of our estimators to test for positive returns to investing in commodities futures contracts. While commodities tend to have positive returns on average, they also have extremely high volatility. We apply our methods to construct the sample mean and standard error of the returns series, and test whether the returns are statistically distinguishable from zero. Due to the structure of commodities futures contracts, the time series of returns have missing observations, and are therefore a natural application for our estimators.

Commodities futures contracts specify the quantity and price of a commodity to be traded on a predetermined expiration date at some point in the future. While some commodities have contracts expiring every month, many have contract months that are irregularly spaced throughout the year. For example, copper futures contracts were available for only March, May, July, September, and December until full monthly coverage began in 1989. Consequently, if we want to calculate the monthly return to investing in commodities futures over a long sample period, our monthly returns data will either reflect a fluctuating time-to-maturity, or will be an incomplete series with irregularly spaced missing observations.

For each commodity, we calculate the return as the percentage change in the price of the contract over the three months prior to the contract's expiration. For instance, we calculate the return to the December contract as the change in price from the last trading day in August to the last trading day in November. Since we want our returns series to reflect the change in the price over the same time-to-maturity, we are only able to calculate this return in the months immediately preceding contract expiration. For copper, this means we will have only five observations each year.[5] The existence of irregularly spaced commodities futures contracts results in a deterministic cyclical pattern of missing observations in the constant-

---

[5]We restrict the full sample for copper to the period with missing contract months, 1960-1989.

maturity returns series. Contract availability and spacing differs across commodities, but tends to remain constant year to year for each commodity.

In this application, we calculate the sample mean for three representative commodities: copper, soybean oil, and lean hogs.[6] We apply our Equal Spacing, Amplitude Modulated, and Imputation estimators to calculate the HAC standard error of the sample mean, and test the hypothesis that the sample mean is significantly different from zero at the five percent level of significance. Under the Imputation method, the missing observations are linearly imputed, while the Equal Spacing and Amplitude Modulated estimators use only the observed data to calculate the mean and standard error. Because it does not provide robust results, we do not provide inference results using the Parzen estimator.

In addition to performing the t-test for the full sample, we use a recursive method to compare our three methods across various sample sizes. For each commodity, we first calculate the sample mean and standard error over the first twelve months of the sample, and perform a t-test of the mean for just this sample window. We then recursively perform the same test for an expanding sample window, adding one month at a time until the full sample is covered. Lastly, we also perform the same type of recursive tests starting with the last twelve months of the full sample. In this backwards recursive test, we use an earlier starting month in each iteration until the sample window again covers the full sample. Having the forward and backwards recursive results allows us to note any structural shifts that may have occurred over time. Figures 2 and 3 depicts the results, and Table 8 provides an overview of rejection rates over the full set of recursive results.

Figure 2 shows the sample mean and 95% confidence intervals constructed using the Imputation and Equal Spacing methods. (The Amplitude Modulated estimator is omitted from the figure for clarity.) The means are very similar across the two methods for most sample windows. The primary difference is that as expected, the Imputation method estimate

---

[6]We selected one representative commodity from each of the major commodity types (metals, animal, and agricultural). We omit the energy commodities, as these commodity contracts do not have missing contract months.

of the standard error is generally smaller than the Equal Spacing and Amplitude Modulated estimates, resulting in a higher rejection rate for Imputation. In the figure, we have shaded the samples for which the hypothesis is rejected under the Imputation method but not rejected under Equal Spacing. The fraction of shaded iterations ranges from 3.2% for lean hogs to 18.2% for copper in the forward recursive results. In the backwards recursive results, the fraction of shaded iterations ranges from 0% for soybean oil to 18.4% for copper.

It is unsurprising that the Imputation method results in a higher rejection rate relative to the Equal Spacing and Amplitude Modulated methods. While in many cases naive imputation is likely to bias the parameter of interest, we have tried to construct an example with little to no bias. However, since the imputed observations are constructed using the observed data rather than drawn from the underlying distribution of data, we note that the standard error of the imputed series is likely lower than the standard error of the observed series. Additionally, the induced high serial correlation of the imputed series will make it likely that the standard error of the imputed series will be underestimated by the Newey-West estimator. For all of these reasons, it is likely that we will have overrejection in hypothesis testing. Without knowing the true mean of the series, in this application we cannot know which method gives us the "right" conclusion for a larger fraction of the tests. Yet, the examples illustrate that the Imputation and Equal Spacing methods can lead to different conclusions in a number of cases, depending on the available data.

7 Conclusion

This paper provides two simple solutions to the common problem of conducting heteroskedasticity and autocorrelation (HAC) robust inference when some observations are missing. Our definitions of the Amplitude Modulated and Equal Spacing estimators are simply formal descriptions of ad hoc practices that are already in use. Yet, by formalizing these procedures, we are able to provide theoretical results that clear up some of the existing confusion

in the literature. By studying the estimators and their properties, we provide justification for their past application to daily business data and through common statistical software packages such as Stata. We also justify their application under a wide variety of missing data structures, including deterministic and cyclical missing structures.

Our theoretical discussion of the estimators highlights a few main conclusions. After establishing the difference between the long-run variance of the underlying and observed series, we demonstrate that our Amplitude Modulated and Equal Spacing estimators both are consistent for the long-run variance of the observed series. This distinction is important, as we also show that we require the long-run variance of the observed series to construct t-statistics for inference, such as in a regression setting. In addition to discussing the asymptotic properties of the estimators, we provide some discussion of their finite sample properties, based on our previous understanding of the finite sample properties of HAC estimators more generally.

We also provide simulation results and apply our estimators to a real world problem involving missing data in commodities futures returns. These results provide further evidence supporting our description of the asymptotic and finite sample behavior of the estimators. In addition, the results of these exercises are used to draw conclusions that can provide guidance to practitioners who need to decide between the estimators for applied work. Though this paper focuses on applying the estimators in a time series setting, they can also be naturally extended for application in a panel setting. We leave this extension for future work.

References

ACHARYA, V. V., AND T. C. JOHNSON (2007): "Insider trading in credit derivatives," *Journal of Financial Economics*, 84(1), 110–141.

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59(3), 817–858.

BEBER, A., M. W. BRANDT, AND K. A. KAVAJECZ (2009): "Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market," *Review of Financial Studies*, 22(3), 925–957.

BLOOMFIELD, P. (1970): "Sepctral Analysis with Randomly Missing Observations," *Journal of the Royal Statistical Society, Series B (Methodological)*, 32(3), 369–380.

BRILLINGER, D. R. (1973): "Estimation of the Mean of a Stationary Time Series by Sampling," *Journal of Applied Probability*, 10(2), 419–431.

———— (1984): "Statistical Inference for Irregularly Observed Processes," in *Time Series Analysis of Irregularly Observed Data*, ed. by E. Parzen, pp. 36–57. Springer-Verlag, New York.

CHOW, G. C., AND A.-L. LIN (1971): "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series," *The Review of Economics and Statistics*, 53(4), 372–375.

CHOW, G. C., AND A.-L. LIN (1976): "Best Linear Unbiased Estimation of Missing Observations in an Economic Time Series," *Journal of the American Statistical Association*, 71(355), 719–721.

CLINGER, W., AND J. W. VAN NESS (1976): "On Unequally Spaced Time Points in Time Series," *The Annals of Statistics*, 4(4), 736–745.

DEN HAAN, W. J., AND A. T. LEVIN (1997): "A Practitioner's Guide to Robust Covariance Matrix Estimation," in *Robust Inference*, ed. by G. Maddala, and C. Rao, vol. 15 of *Handbook of Statistics*, pp. 299–342. Elsevier Science B.V.

DENTON, F. T. (1971): "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization," *Journal of the American Statistical Association*, 66(333), 99–102.

DUNSMUIR, W., AND P. M. ROBINSON (1981a): "Asymptotic Theory for Time Series Containing Missing and Amplitude Modulated Observations," *Sankhyā: The Indian Journal of Statistics, Series A*, 43(3), 260–281.

———— (1981b): "Estimation of Time Series Models in the Presence of Missing Data," *Journal of the American Statistical Association*, 76(375), 560–568.

EATON, J., S. KORTUM, B. NEIMAN, AND J. ROMALIS (2011): "Trade and the Global Recession," *NBER Working Paper No. 16666*.

FORNI, L., L. MONTEFORTE, AND L. SESSA (2009): "The General Equilibrium Effects of Fiscal Policy: Estimates for the Euro Area.," *Journal of Public Economics*, 93(3-4), 559–585.

HARVEY, A. C., AND R. G. PIERSE (1984): "Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, 79(385), 125–131.

IBRAGIMOV, R., AND U. K. MÜLLER (2010): "t-Statistic Based Correlation and Heterogeneity Robust Inference," *Journal of Business and Economic Statistics*, 28(4), 453–468.

JORION, P., AND G. ZHANG (2007): "Good and Bad Credit Contagion: Evidence from Credit Default Swaps," *Journal of Financial Economics*, 84(3), 860–883.

KIEFER, N. M., AND T. J. VOGELSANG (2002a): "Heteroskedasticity-Autocorrelation Robust Standard Errors Using the Bartlett Kernel Without Truncation," *Econometrica*, 70(5), 2093–2095.

——— (2002b): "Heteroskedasticity-Autocorrelation Robust Testing Using Bandwidth Equal to Sample Size," *Econometric Theory*, 18(06), 1350–1366.

——— (2002c): "A New Asymptotic Theory For Heteroskedasticity-Autocorrelation Robust Tests," *Econometric Theory*, 21(06), 1130–1164.

KIEFER, N. M., T. J. VOGELSANG, AND H. BUNZEL (2000): "Simple Robust Testing of Regression Hypotheses," *Econometrica*, 68(3), 695–714.

KIM, M. S., AND Y. SUN (2011): "Spatial Heteroskedasticity and Autocorrelation Consistent Estimation of Covariance Matrix," *Journal of Econometrics*, 160(2), 349–371.

MCKENZIE, C. R., AND C. A. KAPUSCINSKI (1997): "Estimation in a linear model with serially correlated erros when observations are missing," *Mathematics and Computers in Simulation*, 44(1), 1–9.

NEWEY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

——— (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61(4), 631–653.

PAN, J., AND K. J. SINGLETON (2008): "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads," *Journal of Finance*, 63(5), 2345–2384.

PARZEN, E. (1963): "On Spectral Analysis with Missing Observations and Amplitude Modulation," *Sankhyā: The Indian Journal of Statistics, Series A*, 25(4), 383–392.

PETERSEN, M. A. (2009): "Estimating Standard ERrors in Finance Panel Data Sets: Comparing Approaches," *The Review of Financial Studies*, 22(1), 435–480.

PHILLIPS, P. C. B., AND S. N. DURLAUF (1986): "Multiple Time Series Regression with Integrated Processes," *The Review of Economic Studies*, 53(4), 473–495.

PHILLIPS, P. C. B., AND V. SOLO (1992): "Asymptotics for Linear Processes," *The Annals of Statistics*, 20(2), 971–1001.

SCHEINOK, P. A. (1965): "Spectral Analysis with Randomly Missed Observations: The Binomial Case," *The Annals of Mathematical Statistics*, 36(3), 971–977.

WHITE, H. (1984): *Asymptotic Theory for Econometricians*. Academic Press.

ZHANG, B. Y., H. ZHOU, AND H. ZHU (2009): "Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms," *Review of Financial Studies*, 22(12), 5099–5131.

Figure 1: Time Series of Returns
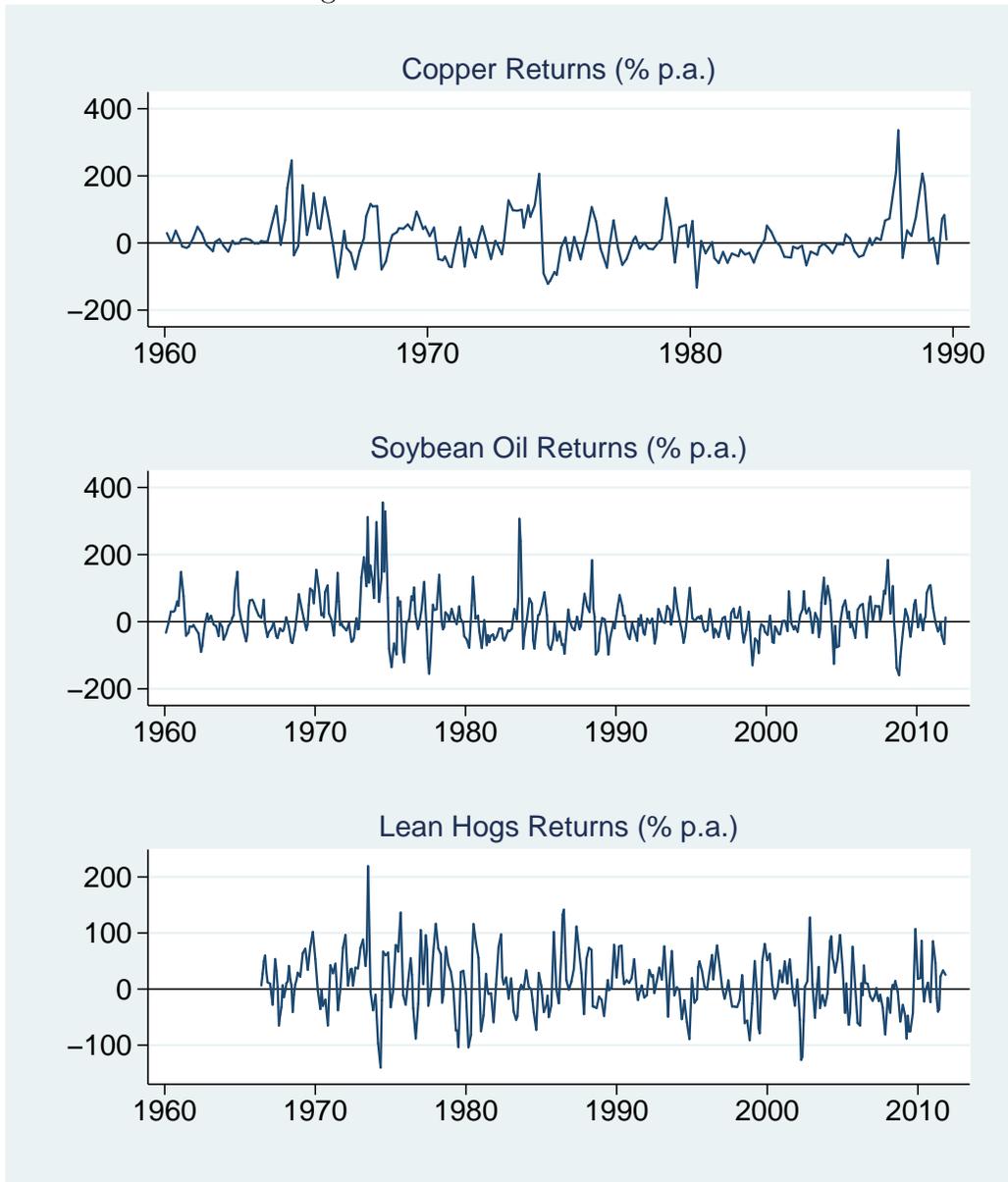
Table 1: Summary of Theoretical Results

Population Parameters	Estimators
Underlying Series:	
$\Omega = \sum_{j=-\infty}^{\infty} \gamma_z(j)$	$\hat{\Omega}^{PZ} \overset{p}{\to} \Omega$
Amplitude Modulated Process:	
$\Omega^* = \sum_{j=-\infty}^{\infty} \kappa(j)\gamma_z(j)$ (Proposition 1)	$\hat{\Omega}^{AM} \overset{p}{\to} \Omega^*$ (Proposition 3)
Equal Spacing Process:	
$\Omega^{ES} = \Omega^*$ (Proposition 4)	$\hat{\Omega}^{ES} \overset{p}{\to} \Omega^*$ (Proposition 5)

Table 2: Benchmark Results

Autocorrelation (ϕ):	Empirical Rejection Rate				
	0.0	0.3	0.5	0.7	0.9
Sample Size T=360					
ES, fixed bw	6.0	7.0	8.0	10.7	23.1
AM, fixed bw	5.6	6.7	8.3	12.6	30.9
Parzen, fixed bw	6.6	4.6	4.5	6.1	19.2
Imputation, fixed bw	8.9	9.7	11.2	15.5	34.4
NW, fixed bw	5.6	7.2	9.1	14.2	33.7
Sample Size T=1200					
ES, fixed bw	5.3	6.0	6.8	8.7	18.8
AM, fixed bw	5.2	6.0	7.2	10.6	26.7
Parzen, fixed bw	5.5	3.6	3.3	4.4	14.9
Imputation, fixed bw	8.1	8.5	9.6	13.1	29.9
NW, fixed bw	5.2	6.4	7.9	11.9	29.4
Sample Size T=24000					
ES, fixed bw	5.1	5.4	5.8	6.6	11.0
AM, fixed bw	5.1	5.4	6.1	7.5	16.5
Parzen, fixed bw	5.1	3.0	2.3	2.2	6.3
Imputation, fixed bw	6.3	6.5	7.0	8.4	17.9
NW, fixed bw	5.1	5.6	6.3	7.9	17.5

Notes: Table reports for a range of sample sizes and autocorrelation parameters the empirical rejection rate as defined in Section 5.2. Data follow our benchmark missing structure, the Bernoulli structure with probability of missing 6/12. The Newey-West estimation provided for comparison uses the full series of simulated data without any missing observations. See Sections 3 and 4 for details on the estimators and the regression context, and Section 5.1 for details on the simulation parameters.

Table 3: Varying Missing Structure

| | Empirical Rejection Rate | | | | |
Autocorrelation (ϕ):	0.0	0.3	0.5	0.7	0.9
Randomly missing, Bernoulli structure					
ES, fixed bw	6.0	7.0	8.0	10.7	23.1
AM, fixed bw	5.6	6.7	8.3	12.6	30.9
Deterministic cyclically missing, All lags observed					
ES, fixed bw	5.9	6.9	7.9	10.1	22.2
AM, fixed bw	5.4	6.5	8.0	11.9	30.1
Deterministic cyclically missing, Some lags unobserved					
ES, fixed bw	6.0	6.5	7.5	9.8	22.1
AM, fixed bw	5.5	6.3	8.1	12.8	32.0
Full sample benchmark					
NW, fixed bw	5.6	7.2	9.1	14.2	33.7

Notes: Table reports for a range of autocorrelation parameters the empirical rejection rate for a sample size of $T = 360$ under varying missing structures as described in Section 5.1: the Bernoulli missing structure, the deterministic cyclical missing structure in which all lags are observed, and the deterministic cyclical missing structure in which some lags are never observed. The probability of missing is $6/12$ for the Bernoulli structure, while the cyclical structures have exactly 6 of 12 observations missing in each cycle. The Newey-West estimation provided for comparison uses the full series of simulated data without any missing observations. See Sections 3 and 4 for details on the estimators and the regression context, and Section 5.1 for details on the simulation parameters.

Table 4: Finite Samples: Fixed and Automatic Bandwidth Selection

	Empirical Rejection Rate					Size-Adjusted Power				
Autocorrelation (ϕ):	0.0	0.3	0.5	0.7	0.9	0.0	0.3	0.5	0.7	0.9
Sample Size T=360										
ES, fixed bw	6.0	7.0	8.0	10.7	23.1	78.9	63.5	49.2	32.7	14.0
AM, fixed bw	5.6	6.7	8.3	12.6	30.9	79.4	64.1	50.0	33.0	14.2
ES, auto bw	6.8	7.7	8.7	10.1	16.8	77.8	62.4	48.2	31.7	13.6
AM, auto bw	6.0	7.0	8.0	9.9	19.7	79.0	63.7	49.1	32.4	13.8
NW, fixed bw	5.6	7.2	9.1	14.2	33.7	97.6	82.3	62.2	38.1	14.7
NW, auto bw	5.9	7.4	8.1	10.0	19.8	97.5	81.5	61.2	37.2	14.4
Sample Size T=1200										
ES, fixed bw	5.3	6.0	6.8	8.7	18.8	80.2	65.2	51.4	34.1	14.5
AM, fixed bw	5.2	6.0	7.2	10.6	26.7	80.4	65.5	51.5	34.3	14.6
ES, auto bw	5.5	6.3	6.9	7.4	10.8	79.9	64.9	50.5	33.5	14.2
AM, auto bw	5.3	6.1	6.4	7.4	12.6	80.2	65.2	51.0	33.9	14.4
NW, fixed bw	5.2	6.4	7.9	11.9	29.4	97.8	83.0	63.0	38.6	15.0
NW, auto bw	5.3	6.3	6.6	7.5	12.6	97.8	82.6	62.7	38.1	14.8
Sample Size T=24000										
ES, fixed bw	5.1	5.4	5.8	6.6	11.0	80.2	65.4	51.0	33.8	14.4
AM, fixed bw	5.1	5.4	6.1	7.5	16.5	80.3	65.4	51.1	33.7	14.4
ES, auto bw	5.1	5.4	5.7	5.7	6.2	80.2	65.3	50.9	33.6	14.4
AM, auto bw	5.1	5.4	5.5	5.7	6.4	80.3	65.4	51.0	33.7	14.3
NW, fixed bw	5.1	5.6	6.3	7.9	17.5	97.8	83.1	63.0	38.6	14.9
NW, auto bw	5.1	5.4	5.5	5.6	6.4	97.8	83.1	63.1	38.7	14.9

Notes: Table reports for a range of sample sizes and autocorrelation parameters the empirical rejection rate and size-adjusted power as defined in Section 5.2. Data follow our benchmark missing structure, the Bernoulli structure with probability of missing 6/12. The Newey-West estimation provided for comparison uses the full series of simulated data without any missing observations. See Sections 3 and 4 for details on the estimators and the regression context, and Section 5.1 for details on the simulation parameters.

Table 5: Empirical Bias and Variance

	Bias					Variance				
Autocorrelation (ϕ):	0.0	0.3	0.5	0.7	0.9	0.0	0.3	0.5	0.7	0.9
Sample Size T=360										
ES, fixed bw	0.00076	0.0201	0.174	3.38	1,051	0.0428	0.103	0.283	1.44	42.5
AM, fixed bw	0.00018	0.0192	0.239	5.69	1,461	0.0273	0.070	0.189	0.86	18.6
ES, auto bw	0.00293	0.0327	0.194	2.22	590	0.0594	0.142	0.449	2.81	132.0
AM, auto bw	0.00072	0.0222	0.154	2.43	819	0.0356	0.098	0.334	2.03	78.9
NW, fixed bw	0.00018	0.0653	0.889	22.01	5,769	0.0220	0.084	0.280	1.55	43.5
NW, auto bw	0.00075	0.0590	0.410	7.70	2,999	0.0309	0.156	0.628	4.93	250.6
Sample Size T=1200										
ES, fixed bw	0.00010	0.0097	0.100	2.19	830	0.0153	0.037	0.105	0.56	18.8
AM, fixed bw	0.00002	0.0111	0.157	4.20	1,277	0.0094	0.024	0.068	0.33	8.1
ES, auto bw	0.00041	0.0127	0.080	0.81	243	0.0215	0.060	0.218	1.34	89.6
AM, auto bw	0.00009	0.0101	0.059	0.92	394	0.0124	0.042	0.144	1.00	55.4
NW, fixed bw	0.00002	0.0414	0.608	16.58	5,085	0.0078	0.0305	0.105	0.61	19.4
NW, auto bw	0.00010	0.0255	0.161	2.93	1,422	0.0109	0.0733	0.282	2.62	179.2
Sample Size T=24000										
ES, fixed bw	0.00000	0.0018	0.021	0.51	308	0.0038	0.011	0.065	3.04	0.0
AM, fixed bw	0.00000	0.0024	0.037	1.18	683	0.0022	0.007	0.037	1.29	0.0
ES, auto bw	0.00000	0.0020	0.011	0.09	18	0.0099	0.039	0.196	20.66	0.0
AM, auto bw	0.00000	0.0017	0.007	0.08	27	0.0069	0.018	0.153	14.94	0.0
NW, fixed bw	0.00000	0.0094	0.146	4.71	2,730	0.0032	0.0119	0.082	3.60	0.0
NW, auto bw	0.00000	0.0034	0.018	0.25	98	0.0117	0.0407	0.451	53.36	0.0

Notes: Table reports empirical bias squared and variance as defined in Section 5.2. Data follow our benchmark missing structure, the Bernoulli structure with probability of missing 6/12. The Newey-West estimation provided for comparison uses the full series of simulated data without any missing observations. See Sections 3 and 4 for details on the estimators and the regression context, and Section 5.1 for details on the simulation parameters.

Table 6: Automatically Selected Bandwidths

	6 of 12 missing				
Autocorrelation (ϕ):	0.0	0.3	0.5	0.7	0.9
T=360, Fixed bandwidth: m(T)=5					
ES, auto bw	5.9	5.1	5.0	6.4	10.0
AM, auto bw	5.4	5.0	6.4	9.8	13.0
NW, auto bw	5.4	5.4	8.0	11.4	14.1
T=1200, Fixed bandwidth: m(T)=6					
ES, auto bw	6.4	6.0	7.0	10.8	17.2
AM, auto bw	6.2	6.7	10.5	16.5	22.2
NW, auto bw	6.2	8.1	13.2	18.8	23.9
T=24000, Fixed bandwidth: m(T)=12					
ES, auto bw	13.2	14.6	20.2	33.3	65.2
AM, auto bw	13.2	18.0	31.8	55.7	94.0
NW, auto bw	13.2	23.5	39.0	62.3	98.3

Notes: Table reports for a range of autocorrelation parameters and sample sizes the mean of the bandwidth selected by the Newey and West (1994) procedure described in Section 5.1. Data follow our benchmark missing structure as described in the Bernoulli structure with probability of missing 6/12. The Newey-West estimation provided for comparison uses the full series of simulated data without any missing observations. See Sections 3 and 4 for details on the estimators and the regression context, and Section 5.1 for details on the simulation parameters.

Table 7: Varying Fraction of Missings

Autocorrelation (ϕ):	Empirical Rejection Rate					Size-Adjusted Power				
	0.0	0.3	0.5	0.7	0.9	0.0	0.3	0.5	0.7	0.9
4 of 12 missing										
ES, fixed bw	5.9	7.0	8.2	11.3	26.6	89.2	72.7	56.2	35.8	14.5
AM, fixed bw	5.6	6.9	8.6	13.1	32.4	89.4	73.2	56.6	36.2	14.6
ES, auto bw	6.4	7.5	8.5	9.7	17.4	88.7	72.0	55.0	34.9	14.0
AM, auto bw	5.9	7.2	8.0	9.8	19.8	89.2	72.5	55.6	35.5	14.2
6 of 12 missing										
ES, fixed bw	6.0	7.0	8.0	10.7	23.1	78.9	63.5	49.2	32.7	14.0
AM, fixed bw	5.6	6.7	8.3	12.6	30.9	79.4	64.1	50.0	33.0	14.2
ES, auto bw	6.8	7.7	8.7	10.1	16.8	77.8	62.4	48.2	31.7	13.6
AM, auto bw	6.0	7.0	8.0	9.9	19.7	79.0	63.7	49.1	32.4	13.8
8 of 12 missing										
ES, fixed bw	6.4	7.1	7.8	9.6	18.7	61.0	50.5	41.2	29.1	13.6
AM, fixed bw	5.6	6.4	7.6	11.2	28.3	62.5	51.9	42.5	29.9	13.9
ES, auto bw	7.8	8.4	9.1	10.5	15.7	57.6	48.2	39.6	27.7	13.1
AM, auto bw	5.9	6.6	7.6	9.5	19.4	61.9	51.4	41.8	29.2	13.6
Full sample benchmark										
NW, fixed bw	5.6	7.2	9.1	14.2	33.7	97.6	82.3	62.2	38.1	14.7
NW, auto bw	5.9	7.4	8.1	10.0	19.8	97.5	81.5	61.2	37.2	14.4

Notes: Table reports for a range of autocorrelation parameters the empirical rejection rate and size-adjusted power for a sample size of $T = 360$ under varying probability of missing observation. Data follow our benchmark missing structure as described in Section 5.1, the Bernoulli structure with probability of missing $4/12$, $6/12$, or $8/12$. The Newey-West estimation provided for comparison uses the full series of simulated data without any missing observations. See Sections 3 and 4 for details on the estimators and the regression context, and Section 5.1 for details on the simulation parameters.

Table 8: Rejection Rates for Recursive Tests

Panel A: 5% significance level

	Copper		Soybean Oil		Lean Hogs	
	Forward	Backward	Forward	Backward	Forward	Backward
Imputation	64.8	28.2	25.3	0.0	93.7	18.3
Amplitude Modulated	59.7	26.5	18.1	0.0	91.4	8.8
Equal Spacing	46.7	9.8	12.4	0.0	90.7	8.4
AM rejects, IM does not	0.0	1.2	0.0	0.0	0.2	0.0
ES rejects, IM does not	0.0	0.0	0.0	0.0	0.2	0.0
IM rejects, AM does not	5.2	2.9	7.2	0.0	2.4	9.5
IM rejects, ES does not	**18.2**	**18.4**	**12.9**	**0.0**	**3.2**	**9.9**

Panel B: 10% significance level

|	Copper		Soybean Oil		Lean Hogs	
	Forward	Backward	Forward	Backward	Forward	Backward
Imputation	77.2	34.6	63.5	22.2	96.1	26.1
Amplitude Modulated	72.9	34.0	29.7	0.0	94.6	15.1
Equal Spacing	65.4	31.4	24.8	0.0	93.5	15.1
AM rejects, IM does not	0.0	0.0	0.0	0.0	0.0	0.0
ES rejects, IM does not	0.0	0.0	0.0	0.0	0.0	0.0
IM rejects, AM does not	4.3	0.6	33.8	22.2	1.5	11.0
IM rejects, ES does not	11.8	3.2	38.7	22.2	2.6	11.0

Notes: Table reports the rejection rate for the test of the hypothesis that the recursive sample mean is equal to zero. The final row of the first panel corresponds to the shaded areas in Figures 2 and 3.

Figure 2: Recursive Test Results - Sample Mean and Error Bands
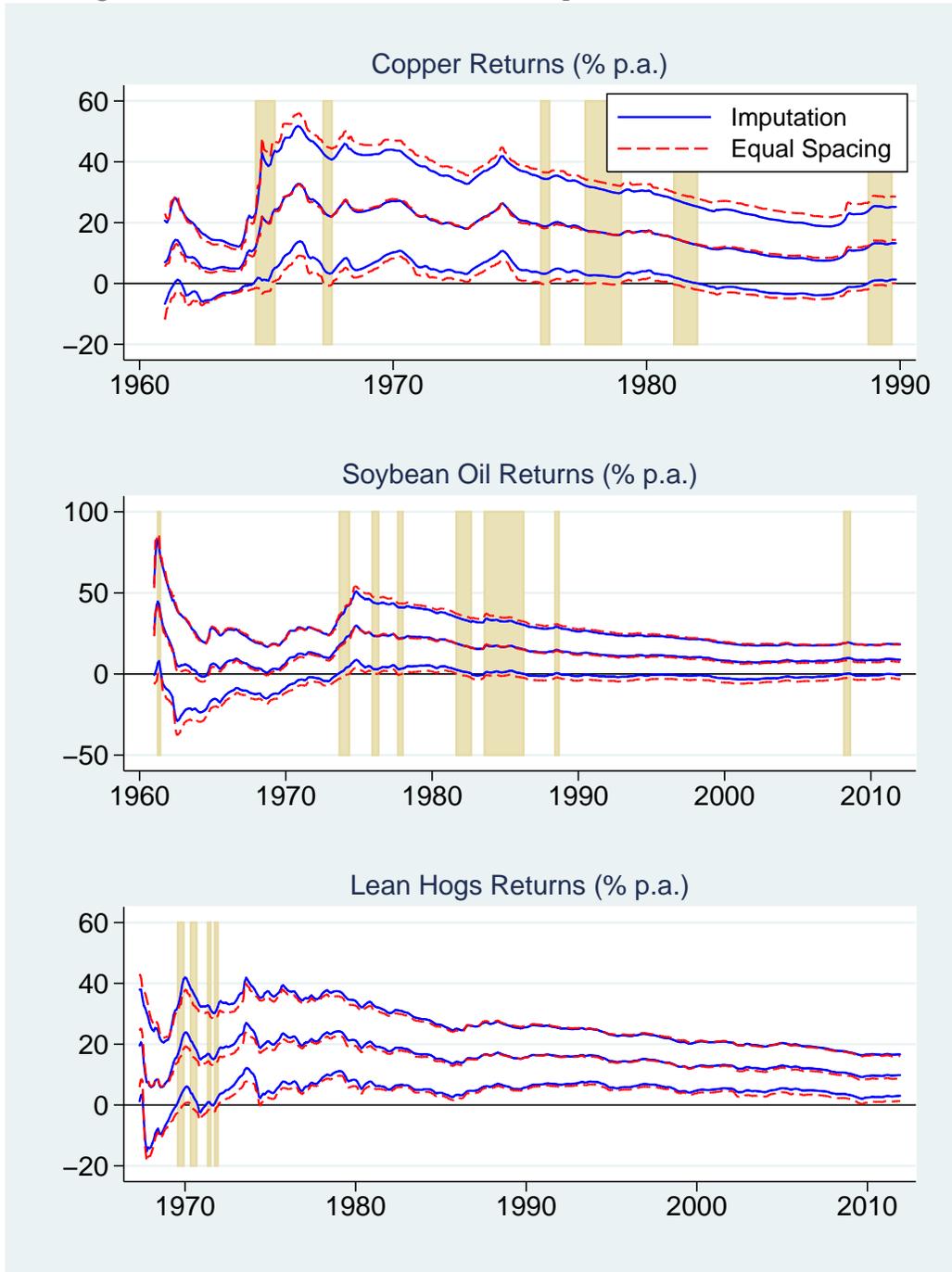
Note: Figure plots the recursive sample mean and 95% confidence intervals for the imputation and equal spacing methods. The recursive sample mean is calculated as the mean of returns over the period from the start of the sample to sample end date (plotted on the x-axis). Shaded areas indicate recursive samples for which the imputation method finds statistical significance while the equal spacing method finds none.

Figure 3: Backwards Recursive Test Results - Sample Mean and Error Bands
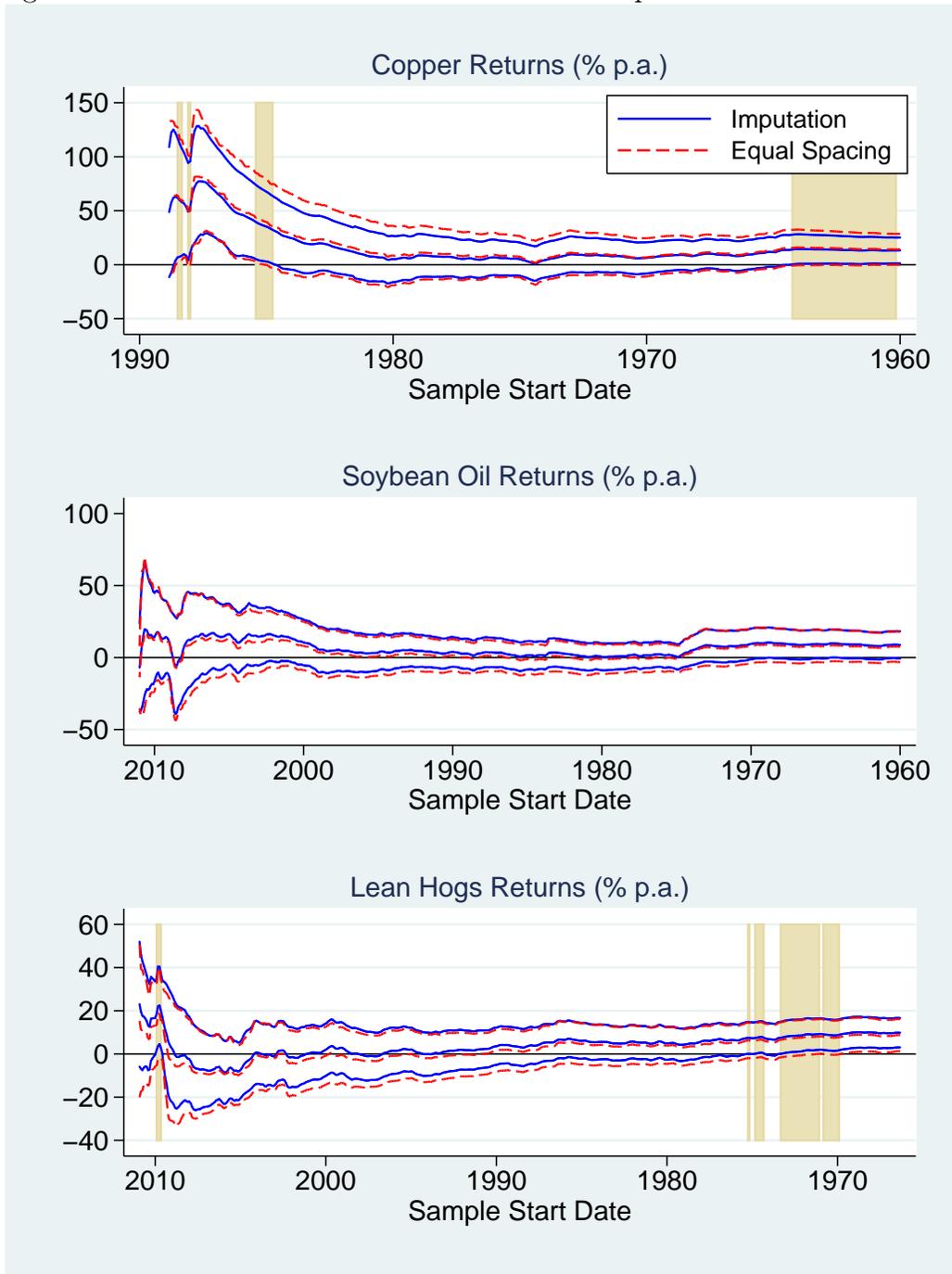
Note: Figure plots the backwards recursive sample mean and 95% confidence intervals for the imputation and equal spacing methods. The backwards recursive sample mean is calculated as the mean of returns over the period from the end of the sample to sample start date (plotted on the x-axis). Note that the sample size is increasing with earlier start dates, moving left to right on the plot. Shaded areas indicate recursive samples for which the imputation method finds statistical significance while the equal spacing method finds none.