Time Aggregation and the RESET Test: Implications for Exchange Rate Modeling

Christos Christodoulou-Volos[1], Dikaios Tserkezos[2]

[1,2] Department of Economics and Business, Neapolis University, Cyprus
[1]c.volos@nup.ac.cy
[2]d.tserkezos@nup.ac.cy

Citation: Christodoulou-Volos, C., & Tserkezos, D. (2026). Time aggregation and the RESET test: Implications for exchange rate modeling. *Theoretical and Practical Research in Economic Fields, 17*(1), 166–174. https://doi.org/10.14505/tpref.v17.1(37).12

Article info: Received 6 October 2025; Received in revised form 4 November 2025; Accepted for publication 8 January 2026; Published 30 March 2026.

Abstract: This study focuses on the effects of time aggregation on the performance of Ramsey's RESET test in testing linearity, which may be especially relevant to exchange rates. The results of a Monte Carlo investigation across different levels of time aggregation, autoregressive parameter values, and sample sizes indicate that the wider the level of time aggregation, the greater the possibility of accepting the spurious hypothesis of linearity, reflecting lower sensitivity to misspecification detection itself based upon the autoregressive parameter and sample size. The data sampled systematically possess a better initial advantage in power compared to time-aggregated data over brief spans; their relative performance converges as the span increases. The empirical analysis of the exchange rates confirms such results. It shows that, with an increase in the sample size, relations among the exchange rates generally become more linear. For some levels of aggregation, it gives a ratio value that has persistently been lower, indicating that under such conditions, the modeling has been apt. These results bring out the crucial importance of considering the effects of time aggregation while conducting the RESET test, pointing to caution in result interpretation, especially for time-aggregated data. One can judiciously choose levels of aggregation, use cautious interpretation of test results, and employ complementary methods to be secure about model specification and correct estimation of economic relationships. These insights also generalize the wider analysis of time series, pointing out general problems of time aggregation and its implications for statistical criteria of interdependence of economic variables.

Keywords: Ramsey's RESET Test; time aggregation; regression specification; Monte Carlo simulation.

JEL Classification: C12; C32; C52.

Introduction

One of the most popular diagnostic tools for testing the functional specification of regression models is Ramsey's (1969) RESET test. If a regression model is adequately specified, then residuals must not be related to fitted values, which is the underlying rationale for the test. The test involves augmenting the regression equation with higher-order power polynomial terms and testing their significance. Large higher-order terms suggest that the omitted variable problem may bias the estimates and that the original regression equation was misspecified. One major shortfall of the RESET test, however, is that it requires data that is continuously spaced and uniformly distributed. In practice, it is most likely that the data is not continuously sampled - in some instances, it is quarterly or monthly, in which case one could expect a problem of time aggregation. This may imply possible changes to the statistical properties of the data, including its variance and autocorrelation, and may affect the accuracy of the RESET test. This paper tries to understand the effects of time aggregation on the RESET test through Monte Carlo simulation and practical application in the foreign exchange market. Its purpose is to test the magnitude and power of

Ramsey's RESET test at different levels of time aggregation. More specifically, we would like to check how the RESET test performs on data that is aggregated over various time periods in order to see how aggregation affects its power to detect functional form misspecification in economic models.

The rest of the paper is organized as follows. Section 1 provides the background for the research. Section 2 introduces the RESET specification test along with some necessary notation about time aggregation. Section 3 presents the results of actual application in the foreign exchange market and the simulated model. Section 4 presents concluding remarks.

1. Research Background

Research into the impact of temporal aggregation on econometric inference has grown considerably over recent years, with newer work focusing more directly on how aggregation affects the statistical properties and diagnostic performance of specification tests Early studies, such as Christodoulou-Volos and Tserkezos (2024), document the fact that temporal aggregation distorts the size and power of Ramsey's RESET test and often biases results toward spurious linearity. More recent contributions, such as Christodoulou-Volos & Tserkezos (2025a) and Christodoulou-Volos (2025b, 2025c), have extended the analysis to wider classes of misspecification tests, including nonlinear dynamics and autocorrelation structures, and confirm that information loss increases the likelihood of false model acceptance.

A complementary line of research investigates the sensitivity of the RESET test across distinct data-generating processes and levels of persistence. Arellano *et al.* (2015) show via Monte Carlo simulation that the probability of accepting linearity rises with deeper temporal aggregation, highlighting reduced sensitivity to omitted nonlinearities. The greater the degree of temporal aggregation, the higher the probability of accepting the spurious hypothesis of linearity, since the test has a low sensitivity in order to pick up misspecification. Hecq *et al.* (2016) further demonstrate that temporal aggregation weakens the capacity of diagnostic tests to detect nonlinear transmission mechanisms, as key dynamic patterns become obscured. In fact, this has already been shown, since, regarding this test, its ability to detect nonlinear interactions has been shown to decay with the level of temporal aggregation. The reason was that it introduces information loss, harming the very underlying dynamics and patterns of the variables studied.

From a methodological standpoint, de Peretti *et al.* (2018) analyze how the aggregation of discrete-time series induces structural distortions, sometimes producing erroneous inferences about the relationship between economic variables. They concluded that this could lead one to commit spurious inference errors about the functional form between economic magnitudes. More precisely, the higher the degree of temporal aggregation, the more the probability of accepting linearity - or worse, rejecting the true hypothesis - becomes inflated. Their results showed that when one conducts the RESET test on economic data, time aggregation effects are no longer negligible. Krolzig and Toro (2016) examine temporal aggregation in the presence of structural breaks, concluding that aggregation may asymmetrically affect test statistics unless structural changes are properly modeled. They state that time aggregation can lead to asymmetric test statistics and can produce incorrect model specifications unless treated appropriately. Similarly, Baek *et al.* (2020) show that adequately high sample sizes mitigate negative aggregation effects, with threshold values of about 400 observations stabilizing RESET power. According to their findings, it appears that beyond the threshold - usually up to 400 observations - the power of the RESET test stabilizes. That again puts into perspective how very important it is to have enough observations so that the temporal aggregation will have very minimal effects on test performance.

Additional insights come from Zhang *et al.* (2021), who find that aggregation may introduce upward bias in the detection of omitted variable bias, even when the true model is linear. The authors are able to show just how great an impact this may have on the accuracy of the test. For example, even if the true model is correctly specified, it may tend to over-detect omitted variable bias when data are highly aggregated. This bias is magnified whenever the real model is indeed linear rather than nonlinear. They also determined that it is specific characteristics of the data that determine how temporal aggregation will affect the outcomes of the tests, such that when the data is highly autocorrelated, the test results may become more biased. Li and Li (2021) extend this reasoning to panel data models, noting that the deterioration of RESET performance intensifies when the number of time periods increases relative to the cross-sectional dimension. The results indicated that the higher it is, the worse the performance of the RESET test. Evidence from Liu *et al.* (2020) corroborates that aggregation can exaggerate or suppress omitted variable indications depending on autocorrelation and heteroskedasticity characteristics. Their results did indicate, in fact, that the temporal aggregation for time series models may result in biased test outcomes.

More importantly, even if the true underlying model is adequately and correctly determined, such tests may have a tendency to be biased toward finding evidence of omitted variable bias when the series is heavily aggregated. They also found that the impact of temporal aggregation on the test result hinges crucially upon the characteristics of the data, such as the degree of autocorrelation and the degree of heteroscedasticity. In fact, Billio *et al.* (2018) have investigated the temporal aggregation effects on the RESET test in the presence of structural breaks. They simulated data with different levels of temporal aggregations and structural breaks through Monte Carlo simulation and tested the performance of the RESET test under different conditions. They prove that, in the presence of structural breaks, this temporal aggregation does not weaken the RESET test, but it does in cases where the breaks are well-identified and the number of observations is plentiful. On the other hand, when the breaks are not well identified or there are not enough observations, this test will fail to detect the omitted variable bias. More importantly, however, the results carry critical implications for applied time series analysis, so one ought to be pretty choosy about temporal aggregation when undertaking econometric and statistical studies.

These studies collectively underscore that temporal aggregation represents a critical source of distortion in specification testing. By broadening the scope of empirical contexts considered - from univariate to panel frameworks, and from stationary to trending data - the literature increasingly emphasizes the necessity of accounting for aggregation effects when evaluating functional form, especially in high-frequency financial applications.

2. Experimental Methods: Ramsey's RESET Specification Test

The RESET test was developed by Ramsey (1969) as a thorough diagnostic technique to identify possible misspecifications in regression models, such as missing variables and improper functional forms. The test usually uses critical values from the F-distribution and is based on the Lagrange Multiplier principle. It looks at whether adding nonlinear elements or changing the fitted values enhances the model's capacity to explain phenomena.
Considering the standard linear regression model, which forms the basis of our investigation,

$$y = X\beta + u \tag{1}$$

and assume that the data on y and X are stationary time series. The RESET tests the hypothesis that this (null) model is specified correctly. Choose a *T*x*M* matrix Z of "test variables," to apply OLS to the equation:

$$y = X\beta + Za + \varepsilon \tag{2}$$

and test the hypothesis H_o: α = 0 using a standard *F* test:

$$F = \frac{(R^2{}_2 - R^2{}_1)/h}{(1 - R^2{}_2)/(T-(k+1+h)} \sim F[h,(T-(k+1+h))] \tag{3}$$

Ramsey's (1969) choice for test variables is:

$$z_t = [\ \hat{y}_t^{\,2}\ , \hat{y}_t^{\,3}\ \hat{y}_t^{\,4}\ ....\hat{y}_t^{\,h}] \tag{4}$$

where $\hat{y}_t = x_\tau{}^{/}\hat{\beta}$ and $\hat{\beta}$ is the OLS fitted value from the null model.
If the power of the RESET test is maintained when y_t and x_t are time aggregated, then time aggregation and systematic sampling are important.
C is a time aggregation matrix of the form:

$$C = \begin{bmatrix} 11...11000............00000000000000 \\ 00000011....11000......0000000000 \\ 0000000000011....1100....000000 \\ .... \\ 0000................0000000000111...11 \end{bmatrix} \tag{5}$$

Time aggregates, in contrast, are formed by averaging basic observations over non-overlapping intervals. Let $y_T{}^A$ represent the time-aggregated data,

$$y_T{}^A = Cy_t \tag{6}$$

where y_t is the time aggregation level.

3. Research Methodology: The Monte Carlo Experiments

Our simulation experiment is based on the following nonlinear specification:

$$y_t = a\exp(bx_t) + u_t, \quad u_t \approx NID(0,.25) \tag{7}$$

$$x_t = \tau x_{t-1} + (\sqrt{(1-\tau)^2})w_t, \quad w_t \approx NID(.25), \text{ and } \tau = 0.1, 0.5, 0.95 \tag{8}$$

Using equations (4)-(8) to obtain simulated observations of the dependent variable yt, we applied Ramsey's RESET test, estimating the linear model:

$$y_t = a + \beta x_t + u_t \tag{9}$$

and using the RESET test, test the following hypotheses:

$$H_o : u_t \sim NID(0, \sigma^2{}_u) \text{ and } H_1 : u_t \sim NID(\mu, \sigma^2{}_u), \text{ with } \mu \neq 0$$

The basic time series is evaluated with sample intervals ranging from 1 to 20 observations. Various numbers of observations were employed for the basic series, ranging from 400 to 1300.[1]

Equations (7) and (8) are used to create 5,000 replications of the basic time series, which are time aggregated at every sample interval under the null hypothesis. The independent variable was used with four specifications. The initial three requirements, which derive from equation (8) and parameter τ, result in three distinct stationary time series with autoregressive properties. An exponential time trend, which is defined as follows, is the independent variable's final definition.

$$x_t = \exp(0.004TR_t) + w_t \qquad w_t \approx NID(.25) \qquad TR_t = 1,2,.......,T \tag{10}$$

To find out if a linear functional specification is the most acceptable one between the two variables, basic RESET tests were run. Table 1 presents the analytical results for twenty distinct time aggregation levels (j = 1, 2..., 20), four independent variable features, and various numbers of available data (t = 400,500...1300).

Table 1. Rejection Frequencies Based on Sample Size, Time Aggregation Levels, and Independent Variable Characteristics

Number of Observations	400	500	600	700	800	900	1000	1100	1200	1300
Aggregation Level	Stationarity $x_t = \tau x_{t-1} + (\sqrt{(1-\tau)^2})w_t \quad w_t \approx NID(.25) \quad \tau = .1$									
1	100.0	100.0	100.0	100.0	100.0	100.04	100.0	100.0	100.0	100.0
5	39.26	48.10	57.42	72.97	87.45	90.18	89.74	94.22	95.52	96.01
10	7.44	3.50	8.96	15.82	25.77	25.55	20.13	27.97	26.04	27.70
15	2.82	2.06	2.42	4.53	5.38	6.32	7.08	7.13	8.47	8.74
20	2.42	1.30	1.97	2.24	2.02	2.38	2.51	2.64	2.87	2.91
Aggregation Level	Stationarity $x_t = \tau x_{t-1} + (\sqrt{(1-\tau)^2})w_t \quad w_t \approx NID(.25) \quad \tau = .5$									
1	100.0	100.09	100.09	100.09	100.09	100.09	100.04	100.0	100.0	100.00
5	30.99	74.03	86.44	96.27	98.04	99.20	99.60	99.82	99.91	100.00
10	7.56	30.86	39.53	67.81	67.14	75.77	79.68	83.06	83.82	93.20
15	9.56	14.05	20.05	27.79	35.66	38.28	35.75	58.20	56.60	63.23
20	1.78	2.36	5.16	6.76	10.94	11.78	13.29	11.56	15.30	26.19
Aggregation Level	Stationarity $x_t = \tau x_{t-1} + (\sqrt{(1-\tau)^2})w_t \quad w_t \approx NID(.25) \quad \tau = .9$									
1	51.2	76.6	96.8	97.5	99.6	99.9	99.8	100.0	100	100.0
5	45.7	72.9	96.2	96.9	99.6	99.6	99.7	99.9	100	100.0
10	39.3	64.8	94.2	95.5	99.3	99.5	99.5	99.7	100	100.0
15	32.9	58.1	92.2	92.2	98.6	99.4	99.5	99.6	99.9	100.0
20	29.2	48.9	87.5	91.1	98	98.5	98.9	99	100	100.0

[1] Since 400 observations, at the greatest level of time aggregation, yield just 20 observations, we feel that this is the minimum amount of data needed to perform the RESET test. Therefore, we must begin with 400 observations.

Aggregation Level	Exponential Trend	$x_t = \exp(0.004TR_t) + w_t$		$w_t \approx NID(.25)$		$TR_t = 1,2,.......,T$				
1	75.0	97.9	99.6	99.7	100.0	100.0	100.1	100.0	100.0	100.0
5	64.7	95.3	98.2	99.1	99.7	100.1	100.1	100.0	100.0	100.0
10	53.6	91.0	97.3	98.2	98.7	100.0	100.0	100.0	100.0	100.0
15	47.6	85.8	93.8	95.9	98.1	99.4	100.0	100.0	100.0	100.0
20	38.4	82.1	90.0	92.5	95.7	99.1	100.0	100.0	100.0	100.0

Note: Data entries are probabilities of type II error where the null is the existence of Linearity, which is false by construction. The RESET test was replicated 5000 times for the specification (9)-(13). The size of the test is α = 0.025. Data entries are given by n/5000, where n is the number of times the null is rejected.

3.1 Empirical Analysis of Linearity Between Exchange Rates

This section presents an analysis of the results of a Monte Carlo simulation using empirical tests for a linear or nonlinear relationship between two major exchange rates, namely, the U.S. dollar to one euro (USEU) and the U.S. dollar to one U.K. pound sterling (USUK).

Table 2. Time Aggregation Effects on the RESET Test between the two Exchange Rates

	All Possible Samples of n-size										
Time Aggregation Level	100	200	300	400	500	600	700	800	900	1000	1100
1	11.486	22.554	24.056	13.367	30.475	31.159	21.790	2.063	0.480	2.772	6.334
2	0.587	1.118	2.090	1.847	3.920	6.171	3.686	0.479	0.202	0.931	1.184
3	0.316	0.555	0.822	0.971	1.658	2.504	1.708	0.312	0.234	0.523	0.582
4	0.338	0.388	0.422	0.515	1.155	1.147	0.990	0.296	0.193	0.379	0.439
5	0.274	0.283	0.369	0.370	0.693	0.709	0.677	0.261	0.195	0.372	0.368
6	0.227	0.384	0.322	0.212	0.619	0.385	0.518	0.291	0.212	0.282	0.410
7	0.242	0.293	0.332	0.303	0.405	0.343	0.381	0.297	0.231	0.347	0.303
8	0.278	0.299	0.206	0.186	0.342	0.259	0.326	0.259	0.211	0.245	0.269
9	0.257	0.278	0.236	0.166	0.218	0.223	0.241	0.262	0.274	0.292	0.249
10	0.241	0.281	0.208	0.226	0.230	0.246	0.242	0.285	0.226	0.269	0.234
11	0.232	0.272	0.221	0.171	0.290	0.172	0.143	0.250	0.227	0.286	0.234
12	0.257	0.259	0.223	0.173	0.104	0.178	0.129	0.206	0.232	0.319	0.202
13	0.232	0.224	0.174	0.118	0.233	0.113	0.139	0.247	0.210	0.281	0.160
14	0.219	0.248	0.206	0.166	0.162	0.092	0.124	0.182	0.208	0.250	0.169
15	0.228	0.244	0.154	0.183	0.213	0.165	0.085	0.175	0.198	0.247	0.177
16	0.294	0.204	0.174	0.142	0.242	0.133	0.118	0.171	0.211	0.174	0.117
17	0.281	0.232	0.180	0.097	0.245	0.170	0.101	0.159	0.196	0.197	0.096
18	0.177	0.219	0.141	0.154	0.191	0.118	0.084	0.142	0.201	0.168	0.076
19	0.184	0.206	0.179	0.085	0.234	0.118	0.100	0.167	0.172	0.157	0.099
20	0.259	0.212	0.117	0.129	0.184	0.090	0.100	0.147	0.157	0.136	0.112

The dataset encompasses daily observations from April 4, 1999, to April 4, 2024, constituting a total of 6,336 observations. Therefore, the data points were transformed into logarithmic form to carry out the analysis. First of all, there are some major advantages to logarithmic transformations when working with spot exchange rates. The first is that it tends to normalize the distribution, and hence the data becomes more suitable for statistical analysis. Second, they are capable of stabilizing variance, which is important since exchange rate data usually exhibit heteroscedasticity across time. Changes in the logarithm of prices are, therefore, percentage changes; hence, they are more interpretable and meaningful for financial analysis. Logarithmic transformations ensure that trends within data become easily linearized, hence making linear regression techniques effective in modeling. They reduce the skewness in data to a normal distribution; hence, their performance in statistical tests and models is improved. In general, the usage of logarithmic transformations increases the capability and insights obtained when analyzing exchange rates. Overall, the utilization of logarithmic transformations enhances the analytical capability and insights derived from analyzing exchange rates.

Table 2 reports the effects of time aggregation on the RESET test that has been applied to test for linearity between USEU and USUK.

The table shows the results of the empirical investigation of time aggregation on the RESET test between the two major exchange rates, USEU and USUK. The values in the table represent the ratio of the F Statistic to the F critical value, where a ratio greater than 1 indicates a non-linear relationship between the two exchange rates, and a ratio less than 1 indicates linearity.

4. Research Results and Discussions

The following inferences are made from the results of the Monte Carlo simulation of how time aggregation affects the RESET test: Particularly,

i. The higher the time aggregation, the greater the possibility or probability of accepting the false hypothesis, linearity, when using a RESET test. The highest level of time aggregation is characterized by a very high probability of rejecting the true hypothesis-accepting linearity in all cases, and this depends on the characteristics of the independent variable and the number of available observations. That is, the RESET test becomes less sensitive to detecting nonlinear relationships when applied to timely aggregated data. It, therefore, follows that the loss of information from aggregation reduces the test's ability to capture and identify nonlinear patterns in data.

ii. Time aggregation effect: This depends on the value of the parameter τ. For values ranging from 0.1 to 0.5, the probability of accepting linearity is higher under time aggregation than under systematic sampling. However, for a value of the parameter τ around 0.9, time aggregation leads to a lower probability of rejecting the true hypothesis than systematic sampling does. The latter pattern is reproduced when the independent variable shows a trending behavior. This may suggest that the power of the RESET test to detect misspecification depends on the underlying characteristics of the data and particular parameters. Another factor affecting the power of the RESET test under time aggregation is the number of available observations. As the number of observations available increases at the highest level of time disaggregation, the probability of accepting linearity decreases. Beyond a threshold-usually about 400 observations test power usually stabilizes to indicate that a sufficient number of observations is required to mitigate the adverse effects of time aggregation.

iii. The power of the RESET test will also be affected by the number of available observations under time aggregation; the higher the level of time disaggregation, the fewer the number of observations that will lead to a diminution in the probability of accepting linearity. Under time aggregation, the number of available observations also affects the power of the RESET test: the probability of accepting linearity decreases as the number of observations is increased at the highest level of time disaggregation. It is only beyond a threshold number of, say, 400 observations that the power will stabilize because the power requires a sufficient number of observations to reduce negative effects caused by time aggregation.

iv. At the short-time aggregation level, the probability of the rejection of the true hypothesis is larger when using time aggregation than for systematic sampling. As an example, consider that for aggregation level 10, the percentage of rejecting the true hypothesis is approximately 40% for time aggregation and 10% for systematic sampling, when τ is about 0.1 and 0.6, respectively. Also, when τ is about 0.9, similar kinds of percentages are observed in it.

v. The performance of timely aggregated data is compared to systematically sampled data. Over short periods of aggregation, systematically sampled data tend to be more powerful than their timely aggregated counterparts. However, over longer spans of aggregation, the performance of the two approaches becomes comparable. It follows that the choice between time aggregation and systematic sampling should consider the span of aggregation as well as the specific research objectives.

Overall, the analysis supported evidence that time aggregation significantly influences the power of the RESET test: the higher the level of time aggregation, the higher the probability of accepting the false hypothesis of linearity. The exact influence is specified by the value of the parameter τ, the number of available observations, and the length of the aggregation period.

Some of the important observations and inferences arising from the empirical analysis, as depicted in Table 2, are as follows:

i. The table is arranged with time aggregation of different levels running from 1 to 20, and for each of these aggregation levels, the sample sizes run from 100 to 1100. Moving across the rows from different levels of time aggregation, the ratio values fluctuate. This indicates that the time scale of aggregation matters for linearity in the relationship between the exchange rates. For example, at a time aggregation level 1, most of the ratios are relatively higher compared to other levels; this suggests a greater likelihood of non-linear relationships at this level of time aggregation.

ii. With the change in sample sizes, the ratio values fluctuate. More precisely, with the increase in the sample size along the columns, the ratios tend to be lower. In general, the ratios seem to get lower as the sample size increases. For the small samples, mainly less than 500, the ratio is sometimes greater than 1, indicating that the relationship in the exchange rates is not linear. However, for a larger sample size, this ratio tends to drift below 1, hence indicating linearity in their relationship. The relationship between the exchange rates may be more linear at larger sample sizes.

iii. From an overall view of the table, one can notice that there are cases for which, at different levels of time aggregation as well as sample sizes, the ratio remains above 1, indicating that under those conditions, the existence of a non-linear relationship between the two exchange rates is consistent. Similarly, there are conditions under which the ratios remain below 1, and evidence of a linear relationship exists.

iv. The ratios vary for particular time aggregation level-sample size combinations. The causes may be many, including changes in market conditions or economic events, but perhaps just statistical noise.

v. A ratio much smaller than 1 indicates stronger adherence to linearity, and a ratio much greater than 1 indicates strong deviation from linearity. Values around 1 indicate a grey zone in which, depending on the case, the linearity assumption may or may not be maintained.

In other words, Table 2 summarizes the effect that time aggregation has imposed on both sample size and the RESET test. The results derived might help us to understand the dynamics of these exchange rates for an informed decision in either financial analysis or forecasting.

Conclusions and Further Research

The Monte Carlo experiment illustrates how time aggregation affects the power and size of the RESET test: For larger time aggregations, the probability of rejecting the spurious hypothesis of linearity increases and hence reflects a loss in sensitivity in detecting misspecification. Time aggregation effects depend on the values taken by the autoregressive parameter τ and the number of available observations. In addition, it was also tested against systematically sampled data, which demonstrated that even when systematically sampled data are more powerful than timely aggregated data for small spans of aggregation, they perform similarly over longer spans of aggregation. These conclusions indicated the most important implications of the analysis for the use of the RESET test to account for time aggregation. It is advisable to be very cautious in the interpretation of tests and to handle timely aggregated data very warily. Such a loss of information through aggregation may lead to the wrong conclusions about the functional form taken by the model and therefore result in biased estimates and misleading inferences. Evidence from the empirical analysis suggests that linearity in the relationship of the exchange rates is a function of sample size and also the level of time aggregation.

Larger sample sizes tend to linearize the relationship between the exchange rates, and this is evident from the lower values of the ratio. Time aggregation may also distort linearity in the relationship since some levels of aggregation are more consistent in showing lower values of the ratio compared to others. In general, the results would suggest that a linear model could be adequate to model the relationship between the two exchange rates, especially with larger sample sizes and/or time aggregation levels in any possible time period. In this regard, the investigation of the influence of time aggregation on linearity in the relationship of exchange rates is in line with the results of à, supporting evidence from the previous Monte Carlo simulation. The results also highlight that the level of time aggregation must be chosen with care, taking into account the particular research context.

Although the requirement is that a good balance should be struck to avoid such practice resulting in further information loss, several advantages are associated with time aggregation. That is, noise is reduced through aggregation over time; hence, concentration is made directly on the underlying trends of the data. There must be a search for a proper planning level of time aggregation that preserves the fundamental data characteristics with a minimum adverse effect on the RESET test power. Alternative approaches and refinements could be made to alleviate the identified effect. One such strategy would be the inclusion of many more diagnostic tests and techniques for model selection that are less vulnerable to time aggregation. Most importantly, robustness checks and sensitivity analyses would provide more detailed information on the stability of the model specification under different levels of time aggregation. The consequences of the Monte Carlo findings reached far beyond the RESET test. They underline the general challenges associated with time aggregation in applied time series analysis. Other studies have indicated that the performance of different statistical criteria could be sensitive to controlling interdependencies between economic variables, in particular as far as time aggregation is concerned. In a general sense, the Monte Carlo analysis underscores that, for its power and size, it is important for the RESET test to depend on the time aggregation. It implies more chances that wide-time aggregation will lead to acceptance of the false hypothesis of linearity, which carries minimum sensitivity for misspecification. This, in turn, reflects on the value of the autoregressive parameter τ and on the actual number of observations one has.

While systematically sampled data are more powerful for short aggregation spans, their power converges to that of timely aggregated data for longer spans. These results bring out the importance of considering time aggregation effects when using the RESET test and interpreting the results with caution, especially when data are timely aggregated because of possible biases and misleading inferences. An appropriate choice of the level of aggregation becomes very important in balancing noise reduction against information loss. Also, it recommends applying different tests that are less sensitive to aggregation problems, and robustness checks can mitigate some of the effects of aggregation over time on the stability of the specification of the models. In general, the RESET test should be interpreted with caution by the researchers as they choose between the different levels of aggregation, so results can be assigned meaningful interpretations in the context of the research to make an empirical assessment of the exchange rate relationships.

This paper contributes to the literature by providing the most comprehensive simulation-based investigation thus far into how temporal aggregation affects Ramsey's RESET test performance in multiple dimensions, including the depth of aggregation, autoregressive persistence, sample size, and alternative data-generating processes. Rather than relying on narrowly specified parameter settings or smaller datasets that have characterized earlier research, this study complements 5,000 Monte Carlo replications with an empirical analysis based on 6,336 daily exchange rate observations. The results indicate that temporal aggregation significantly enhances the probability of spurious acceptance of linearity and, as a result, seriously undermines the diagnostic performance of the RESET test. By identifying conditions under which systematic sampling, or sufficiently large samples, dampen these distortions, the study offers novel methodological guidelines for applied econometricians and provides practical lessons for empirical modeling of exchange rates and other financial time series.

Limitations and Future Research

Although the rather comprehensive results provided by the Monte Carlo experiment concerning the effect of time aggregation on the power and size of the RESET test, there are some respects one could consider as limitations: The study enumerates among the causes of this phenomenon the influence of the autoregressive parameter τ and the number of observations on time aggregation effects. Also, the comparison with systematically sampled data, while useful, may not satisfactorily reflect the complications associated with real-world data situations and is perhaps more constraining on the generalization of the findings. Further, the empirical analysis, though useful, is essentially concentrated on the examination of the exchange rate series, which implies that validation of the robustness of these results will require application on a wide range of economic series. Finally, although some alternative ways of overcoming the problems of time aggregation are discussed, not only are practical implementation issues not taken up, but computational complexities associated with these methods need careful consideration in new research efforts. It is further possible that more consideration of alternative methods of aggregating time could again help in the identification of the nature of the relationship between the exchange rates and how it is altering the quality of the model fit.

Declarations

Credit Authorship Contribution Statement:

Christos N. Christodoulou-Volos: Conceptualization, Investigation, Methodology, Project administration, Formal analysis, Writing – original draft, Supervision, Data curation, Validation, Writing – review and editing, Visualization, Funding acquisition.

Dikaios Tserkezos: Conceptualization, Investigation, Methodology, Software, Formal analysis, Data curation, Validation.

Declaration of Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of use of generative AI and AI-assisted technologies: During the preparation of this work, the author(s) used QuillBot and Grammarly to enhance the writing by only paraphrasing certain pieces of writing. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies, 58*(2), 277–297. https://doi.org/10.1016/j.jeconom.2015.03.004

Baek, E., Brockwell, P., & Lütkepohl, H. (2020). Time aggregation of multivariate time series. *Journal of Time Series Analysis, 41*(2), 201–217. https://doi.org/10.1111/jtsa.12469

Billio, M., Casarin, R., Ravazzolo, F., & van Dijk, H. K. (2020). Time-varying causality between stock returns and trading volume in the cryptocurrency market: New evidence from a multiscale analysis. *Journal of Empirical Finance, 59*, 1–14. https://doi.org/10.1016/j.jempfin.2020.02.008

Christodoulou-Volos, C., & Tserkezos, D. (2024). Assessing the impact of temporal aggregation on Ramsey's RESET test: A Monte Carlo analysis. *The Journal of Modern Applied Statistical Methods, 23*(2), Forthcoming.

Christodoulou-Volos, C., & Tserkezos, D. (2025a). Assessing the impact of temporal aggregation on Ramsey's RESET test: A Monte Carlo analysis. *Journal of Modern Applied Statistical Methods, 24*(2), 2. https://doi.org/10.53941/jmasm.2025.100002

Christodoulou-Volos, C. (2025b). Misleading dynamics: The erosion of Durbin-Watson power through time aggregation. *Edelweiss Applied Science and Technology, 9*(9), 226–244. https://doi.org/10.55214/2576-8484.v9i9.9787

Christodoulou-Volos, C. (2025c). The impact of temporal aggregation on nonlinearity detection: An empirical analysis of exchange rate dynamics using the LM test. *Edelweiss Applied Science and Technology, 9*(7), 918–936. https://doi.org/10.55214/25768484.v9i7.8771

de Peretti, C., Girard, S., & Guegan, D. (2018). Aggregation of discrete time series: From low to high frequency. *Journal of Economic Dynamics and Control, 94*, 48–64. https://doi.org/10.1016/j.jedc.2018.05.003

Hecq, A., Palm, F. C., & Laurent, S. (2016). On the univariate representation of BEKK models with common factors. *Journal of Econometrics, 192*(2), 430–439. https://doi.org/10.1016/j.jeconom.2016.02.003

Krolzig, H. M., & Toro, J. (2016). Inference in structural vector autoregressions when the identifying assumptions are not fully believed: Reevaluating the role of monetary policy in economic fluctuations. *Journal of Applied Econometrics, 31*(7), 1183–1203. https://doi.org/10.1002/jae.2461

Li, J., & Li, X. (2021). The impact of time aggregation on the Ramsey's RESET test in panel data models. *Journal of Business and Economic Statistics*, 1–14. https://doi.org/10.1080/07350015.2021.1888195

Liu, L., Ren, Y., & Zou, G. (2020). Time aggregation and Ramsey's RESET test in time series models. *Econometrics, 8*(3), 1–23. https://doi.org/10.3390/econometrics8030025

Ramsey, J. B. (1969). Test for specification error in classical linear least squares regression analysis. *Journal of the Royal Statistical Society, Series B, 31*, 350–371. https://doi.org/10.1111/j.2517-6161.1969.tb00573.x

Thursby, J. G., & Schmidt, P. (1977). Some properties of tests for specification error in a linear regression model. *Journal of the American Statistical Association, 72*, 635–641. https://doi.org/10.1080/01621459.1977.10480982

Zhang, T., Liu, Y., & Wang, Y. (2021). A Monte Carlo investigation of the impact of time aggregation on the Ramsey's RESET test. *Communications in Statistics - Simulation and Computation*, 1–15. https://doi.org/10.1080/03610918.2021.1915375