Kunst, Robert M.; Franses, Philip Hans

Working Paper

Asymmetric time aggregation and its potential benefits for forecasting annual data

Reihe Ökonomie / Economics Series, No. 252

Provided in Cooperation with:
Institute for Advanced Studies (IHS), Vienna

Suggested Citation: Kunst, Robert M.; Franses, Philip Hans (2010) : Asymmetric time aggregation and its potential benefits for forecasting annual data, Reihe Ökonomie / Economics Series, No. 252, Institute for Advanced Studies (IHS), Vienna

This Version is available at:
https://hdl.handle.net/10419/72726

252

Reihe Ökonomie

Economics Series

Asymmetric Time Aggregation and its Potential Benefits for Forecasting Annual Data

Robert M. Kunst, Philip Hans Franses

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

252

Reihe Ökonomie

Economics Series

Asymmetric Time Aggregation and its Potential Benefits for Forecasting Annual Data

Robert M. Kunst, Philip Hans Franses

July 2010

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Robert M. Kunst
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
☎: +43/1/599 91-255
email: kunst@ihs.ac.at
and
Department of Economics
University of Vienna

Philip Hans Franses
Erasmus School of Economics
Econometrics
Erasmus University Rotterdam
3000 DR Rotterdam, The Netherlands
email: franses@ese.eur.nl

Founded in 1963 by two prominent Austrians living in exile – the sociologist Paul F. Lazarsfeld and the economist Oskar Morgenstern – with the financial support from the Ford Foundation, the Austrian Federal Ministry of Education and the City of Vienna, the Institute for Advanced Studies (IHS) is the first institution for postgraduate education and research in economics and the social sciences in Austria. The **Economics Series** presents research done at the Department of Economics and Finance and aims to share "work in progress" in a timely way before formal publication. As usual, authors bear full responsibility for the content of their contributions.

Das Institut für Höhere Studien (IHS) wurde im Jahr 1963 von zwei prominenten Exilösterreichern – dem Soziologen Paul F. Lazarsfeld und dem Ökonomen Oskar Morgenstern – mit Hilfe der Ford-Stiftung, des Österreichischen Bundesministeriums für Unterricht und der Stadt Wien gegründet und ist somit die erste nachuniversitäre Lehr- und Forschungsstätte für die Sozial- und Wirtschafts-wissenschaften in Österreich. Die **Reihe Ökonomie** bietet Einblick in die Forschungsarbeit der Abteilung für Ökonomie und Finanzwirtschaft und verfolgt das Ziel, abteilungsinterne Diskussionsbeiträge einer breiteren fachinternen Öffentlichkeit zugänglich zu machen. Die inhaltliche Verantwortung für die veröffentlichten Beiträge liegt bei den Autoren und Autorinnen.

Abstract

For many economic time-series variables that are observed regularly and frequently, for example weekly, the underlying activity is not distributed uniformly across the year. For the aim of predicting annual data, one may consider temporal aggregation into larger subannual units based on an activity time scale instead of calendar time. Such a scheme may strike a balance between annual modelling (which processes little information) and modelling at the finest available frequency (which may lead to an excessive parameter dimension), and it may also outperform modelling calendar time units (with some months or quarters containing more information than others). We suggest an algorithm that performs an approximate inversion of the inherent seasonal time deformation. We illustrate the procedure using weekly data for temporary staffing services.

Keywords

Seasonality, time deformation, prediction, time series

JEL Classification

C22, C53

Contents

1 Introduction

We are interested in predicting time-series variables that are available at a subannual frequency. For example, this frequency—in the following called the 'fine' frequency—could be weeks. The objective is to predict the annual values, and we generally assume that the variable is a flow, such that the annual value is the cumulated sum of all weekly values for that year. If the variable is a stock, this does not change the main arguments, as long as the average over all weeks is in focus, as that annual value is just a multiple of the sum. In fact, the empirical example that we present here corresponds to such a stock case.

Obvious suggestions are to forecast the annual variable on the basis of an annual model or of a model tuned to the fine frequency. It is well known that prediction based on the subannual data and on subsequent time aggregation of the multi-step predictions is not always optimal (see, e.g., Man, 2004). Small samples and limited degrees of freedom can entail very inefficient forecasts based on models for the fine frequency. Thus, another alternative could be a partial time aggregation to a 'crude' subannual frequency and to consider time-series modelling on that frequency. In the example of weekly data availability, the crude frequency could be months or quarters.

Quite often, however, economic activity is concentrated in specific parts of the year. For example, tourism in a holiday resort may be low in November and booming around Christmas and in summer. Then, much information will be contained in the third and fourth quarters and little information in the second quarter.

A forecaster may consider forming pseudo-quarters, aggregating the first four months into one observation instead of the first three. This is what will be called regrouping in the following.

We consider an algorithm that aims at spreading the seasonal variation approximately uniformly across the crude frequency. We investigate cases where forecasts based on such an artificial crude data outperform those based on natural splits and sometimes even those based on the fine frequency. This procedure could be classified as a *type-III aggregation* procedure in the terminology of Jordà and Marcellino (2004) that aggregates regularly spaced time points into irregularly spaced time points, albeit with the ultimate aim of predicting a regularly spaced series. Thus, the type-III aggregation is followed by a type-II aggregation that aggregates irregularly spaced data to a final regularly spaced annual series.

Our generating model at the fine frequency builds on the concept of time deformation. Time deformation was introduced to the econometric literature by Clark (1973) whose ideas were followed extensively in finance. Stock (1987, 1988) used the concept for business cycles at a longer and slightly irregular frequency, inspired by the historical work by Burns and Mitchell (1946). By contrast, we use time deformation to describe seasonal behavior within a year. The economic clock is assumed to run faster in certain seasons and to slow down afterwards. The algorithm approximates a reversion of the underlying time deformation. We demonstrate that the success of this procedure depends on the specific design.

We illustrate the procedure using weekly data on temporary staffing services.

The plan of this paper is as follows. Section 2 reviews some simple cases of underlying time-series generating laws and evaluates the properties of annual and subannual forecasting schemes for each example. Section 3 focuses on the concept of time deformation and introduces the time-deformation functions that will be used in the following section. Section 4 introduces our algorithm that targets an approximate inversion of the underlying time deformation in given data. Section 5 reports some simulation experiments for time-deformed data. Section 6 analyzes an empirical application. Section 7 concludes.

2 Some role-model examples

In this section, we analyze some basic seasonal generating processes and their implications with regard to the prospects of regrouping on prediction. These examples are traditional in the sense that they do not refer to the concept of time deformation used in subsequent sections. We convene that the subscript τ denotes the year and w denotes the season within the year $w = 1, \ldots, S$. Double subscripts τ, w denote season w in year τ, with the convention that $X_{\tau,1}$ is preceded by $X_{\tau-1,S}$, for example. Details on calculation are deferred to the appendix.

Example 1. Assume $Z_\tau = \sum_{w=1}^{S} X_{\tau,w}$, where $(X_{\tau,w})$ is a random walk such that $X_{\tau,w} = X_{\tau,w-1} + \varepsilon_{\tau,w}$. $(\varepsilon_{\tau,w})$ is independent white noise with variance σ_ε^2. Then, as shown in the appendix, the

conditional expectation forecast formed at time point (τ, S) for $Z_{\tau+1}$ will be $SX_{\tau,S}$ when it is based on the subannual information. The mean squared error of the subannual forecast is $\sigma_\varepsilon^2 \sum_{j=1}^{S} j^2$. If only annual information is available, $Z_\tau - Z_{\tau-1}$ follows a first-order MA process. The MSE of the naive annual forecast Z_τ is $\sigma_\varepsilon^2(\sum_{j=1}^{S} j^2 + \sum_{j=1}^{S-1} j^2)$, which can exceed the subannual MSE substantially for larger S. The optimal annual forecast yields an only slightly smaller MSE, typically still much in excess of the subannual MSE. Now suppose we can regroup into two groups. Then, the optimal prediction grouping will consist of $\sum_{w=1}^{S-1} X_{\tau,w}$ in the first group and $X_{\tau,S}$ in the second group. The last season is isolated, as it contains the most recent information that is most useful for predicting the upcoming year. The forecast based on the last season only corresponds exactly to the forecast based on the full subannual information.

Example 2. Assume $X_{\tau,w} = X_{\tau-1,w} + \varepsilon_{\tau,w}$, a seasonal random walk, otherwise keep the design of example 1. Then, the forecast based on annual data and the forecast based on subannual data are identical, such that seasonal information has no additional value. This result is invariant to any regrouping of the seasonals.

Example 3. Assume $X_{\tau,w} = \delta_w + \varepsilon_{\tau,w}$, purely deterministic seasonality. Then, the conditional expectation forecasts based on annual as well as subannual data are identically $\sum_{w=1}^{S} \delta_w$, which yields a forecast variance of $S\sigma_\varepsilon^2$. Again, the result is invariant to any regrouping.

Example 4. Assume $S = 4$, and $X_{\tau,w}$ is generated by a periodic process that is a seasonal random walk for $w = 1, 2$

and white noise for $w = 3, 4$. Quarterly data entail a MSE of $4\sigma_\varepsilon^2$. If only annual data are available, the MSE increases to around $5.236\sigma_\varepsilon^2$, as derived in the appendix. Forecasts based on regrouping fall in between these two benchmarks. Traditional semesters, with the first semester formed from quarters 1 and 2, yield a seasonal random walk for the first semester and white noise for the second semester. The MSE is the same as in the quarterly case. Grouping the first quarter separately from the three other quarters yields a seasonal random walk for the first quarter and an ARIMA(0,1,1) for the second group formed from quarters 2–4. The resulting forecast has a variance of $5\sigma_\varepsilon^2$, closer to the inefficient annual than to the efficient quarterly.

Example 5. Modify the conditions in Example 4 such that the seasonal random walk in the first two quarters is replaced by a random walk that adds a white-noise term to the first quarter for the second quarter and another white-noise term to the second quarter for the first quarter of the next year. In this design, the optimal forecast for the next year based on quarters just uses twice $X_{\tau-1,2}$ and achieves a forecast error variance of $7\sigma_\varepsilon^2$. By contrast, for the forecast based on annual data error variance increases palpably to $10\sigma_\varepsilon^2$. Here, grouping plays a role. Direct semester grouping yields an ARIMA(0,1,1) model with the implied forecast error of $7.84\sigma_\varepsilon^2$. By contrast, splitting the year into a first quarter and the remainder yields a forecast that is difficult to evaluate analytically, due to the complex correlation structure among forecast errors. We used some Monte Carlo that indeed conformed to the analytical results for the other cases. For the

regrouped prediction, it yields an error variance well above $9\sigma_\varepsilon^2$. The intuitively beneficial regrouping entails an inefficient forecast that is closer to the uninformative annual than to the most informative quarterly prediction.

Clearly, the reaction to regrouping the subannual observations is quite heterogeneous across data-generating processes. While no general recommendation can be given based on these examples, it appears that strong and homoskedastic seasonality usually entails robustness to regrouping, while strong serial correlation with weak seasonal cycles may lead to considerable sensitivity to regrouping. Collecting subannual observations with similar time-series dynamics can be beneficial for regrouping schemes, and seasonal heteroskedasticity can also be influential.

3 Time deformation

In a sense, the concept of time deformation is ubiquitous in observed economic variables. Economic activity on stock markets, for example, is low while the stock market is closed, such that it may give the impression that time is running faster while the stock exchange is operating. Similarly, heating is less needed in summer, such that the economic time of heating is running faster in winter, slowing down in spring and coming to a standstill on a hot day.

Following a seminal contribution by Clark (1973), the concept has been applied often to financial data with high observational frequency (for example, see Ghysels *et al.*, 1995), and

less often to business cycles (see Stock, 1987, 1988) or to intra-year seasonal variation (Jordà and Marcellino, 2004). For a theoretical exposition of results on a class of time-deformation models, see also Jiang *et al.* (2006).

Seasonal variation, however, has the advantage that the starting points and ends of the intervals of concern are exogenous, at the begin and end of calendar years. By contrast, limits of business cycles are more difficult to determine, and the definition of troughs and peaks may be subject to some subjective expertise, such as the known NBER chronology.

Within the time interval of concern—in the case of seasonality, one year—a deformation function $s = g(t)$ defines the transformation of calendar time t to economic time s. Typically, the function should be invertible, continuous, and monotonic. It may be acceptable to admit violations of strict monotonicity and to allow for episodes without any economic activity.

An example of a deformation function is plotted in Figure 1. The graph depicts the function

$$s = t^{0.2}$$

on the unit interval. Dotted lines signal the points where one, two, and three quarters of calendar time have elapsed. The first quarter is the most active, and actually almost 80% of economic activity is located in the first three months of the year, if the unit interval is equated to a calendar year.

A comparable plot for the time deformation function

$$s = \frac{\arcsin t}{\pi/2}$$

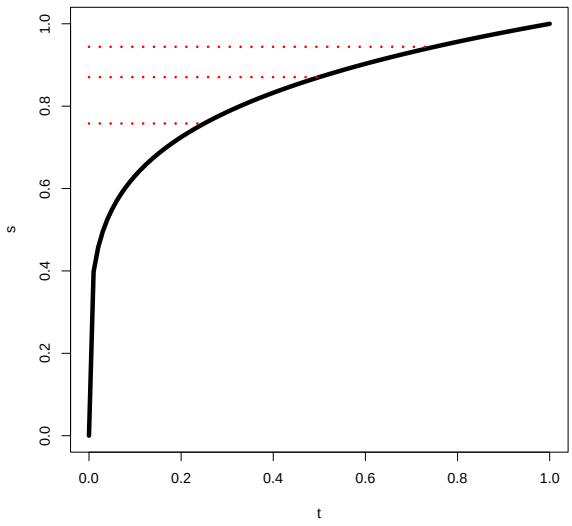

Figure 1: The deformation function $t^{0.2}$ in $[0, 1]$.

is given as Figure 2. Economic activity is assumed to run faster toward the end of the year rather than at the beginning. Generally, time deformation appears to be less radical than for Figure 1.

These two deformation functions will serve as role models for the generating processes in the simulations. It is easy to generalize these functions to allow for several high-activity episodes through the year but we wish to keep the design as simple as possible.

4 The algorithm

Assume a time-series variable $X_{\tau,w}$ is observed at a subannual 'fine' frequency S_1, such that observations $X_{\tau,w}$ are available for

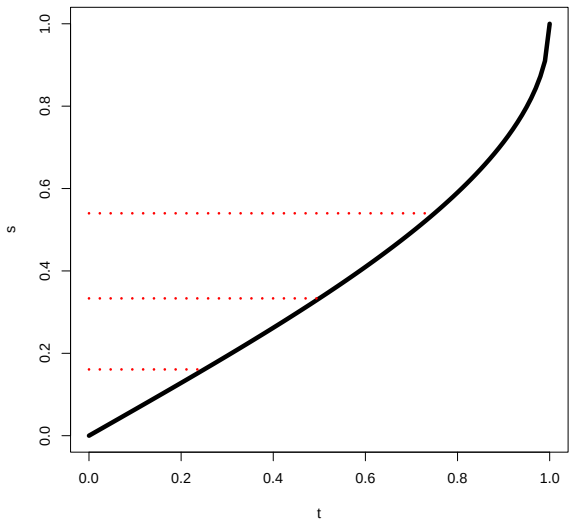

Figure 2: The deformation function $\frac{\arcsin t}{\pi/2}$ in $[0, 1]$.

$\tau = 1, \ldots, T$ and $w = 1, \ldots, S_1$. An annual variable Z_τ is defined as a time aggregate:

$$Z_\tau = \sum_{w=1}^{S_1} X_{\tau,w}, \tau = 1, \ldots, T.$$

For example, $S_1 = 52$ represents weekly availability and $S_1 = 12$ corresponds to monthly data.

The objective is to forecast the next annual value Z_{T+1}, i.e. to find a function of the observed values that approximates Z_{T+1} as closely as possible. Considered criteria for prediction accuracy are the mean squared error (MSE), the mean absolute error (MAE), and the relative ranking among rival predictors.

A traditional approach for model-based prediction is modelling on an annual basis:

$$Z_\tau = f(Z_{\tau-1}, Z_{\tau-2}, \ldots) + \varepsilon_\tau,$$

usually with a linear function $f(.)$, and plugging in the estimated structure to approximate conditional expectations:

$$\hat{Z}_{\tau,I} = \hat{f}(Z_{\tau-1}, Z_{\tau-2}, \ldots).$$

This forecast will be called forecast I in the following, and we will focus on autoregressive f models with a lag order determined by information criteria.

Alternatively, modelling may rely on the S_1 frequency, using x_w in short for $x_{\tau,w}$:

$$x_w = f(x_{w-1}, x_{w-2}, \ldots) + \varepsilon_w.$$

In order to capture the seasonal variation in the data, the f function may contain seasonal dummy variables for the S_1 seasons. In practice, with unknown parameters, this approach requires the estimation of S_1 parameters on top of the p coefficients of an autoregressive specification. Once the model has been estimated from data, forecasts at horizons 1 to S_1 can be generated by iteratively using

$$\hat{x}_w = \hat{f}(x_{w-1}, x_{w-2}, \ldots; \delta_w),$$

where δ_w denotes the seasonal dummy variables. This first yields $\hat{x}_{T+1,1}$. At horizon 2, the unobserved first argument $x_{T+1,1}$ is replaced by the one-step forecast, and this scheme is followed until $\hat{x}_{T,S_1}$ is attained. Finally, time aggregation yields

$$\hat{Z}_{\tau,II} = \sum_{w=1}^{S_1} \hat{x}_{\tau,w}.$$

This forecast will be called forecast II in the following.

An intermediate approach may rely on a 'crude' frequency S_2 that is assumed to be an integer factor of S_1. For example, if $S_1 = 52$, then $S_2 = 4$ or $S_2 = 2$ are possible values. $S_1 = 12$ would admit $S_2 = 2, 3, 4$. Using S_2 instead of S_1 may be motivated by the promise of a simpler model structure and thus more efficient forecasts. We shall see that indeed forecasts based on slightly cruder frequencies tend to dominate those based on extremely fine frequencies.

The traditional approach is to aggregate the data at frequency S_1 to the cruder frequency S_2 by keeping equidistant time points. Then, a modelling technique analogous to forecast II results in a forecast III. This requires the estimation of $S_2 < S_1$ dummy coefficients, and the resulting gain in degrees of freedom can be used to extend the lag order of the autoregression.

Finally, the regrouping approach proceeds as follows. Like model III, it is also based on the frequency S_2. However, it starts from seasonal variance estimates

$$\hat{\sigma}_w^2 = \frac{1}{T-1} \sum_{\tau=1}^{T} (x_{\tau,w} - \bar{x}_w)^2, w = 1, \ldots, S_1.$$

The sum of these variance estimates $\hat{\sigma}^2$ measures the dispersion in the observed variable, although of course it is not tantamount to the straightforward sample variance. For ease of notation, we will not use hats in the following. We now aim at distributing this variation σ^2 uniformly across the year. This is done as follows. The first artificial observation $x_{\tau;1}$ accumulates all original data points $x_{\tau,w}, w = 1, \ldots, K$, such that

$$K = \min\{k : \sum_{t=1}^{k-1} \sigma_w^2 + \frac{1}{2}\sigma_k^2 > \frac{\sigma^2}{S_2}.\}$$

The observation $x_{\tau,K+1}$ at frequency S_1 starts the second artificial observation $x_{\tau;2}$. This scheme is followed in analogy until all observations are regrouped, for all $\tau = 1, \ldots, T$. The forecast based on this algorithm will be called forecast IV in the following.

The intuition behind our regrouping scheme and, in particular, our definition of K is simple. The basic aim is to distribute the variance contributions uniformly across the crude-frequency units. As long as the sum of the fine-frequency variance contributions is less than the proportional share, fine-frequency data are accumulated. When the accumulated sum of variances transgresses this proportional share, the marginal component is allotted either to the 'old' regrouped data point or a new one is started, depending on whether its half value is less or larger than the remaining discrepancy.

A difference in calculating forecast IV relative to forecast II and III is that no seasonal dummy constants are used in constructing the time-series models. We experimented with adding dummy constants but this led to a deterioration of predictive accuracy in all experiments. This is also well in line with the original intention of regrouping as a crude means of reverting the time deformation in the data, which by construction should result in a non-seasonal time series.

5 Monte Carlo evidence

As Section 2 demonstrates, the reaction of forecast precision to time aggregation can be quite heterogeneous. The simulation ex-

periments reported in this section provide additional insight into such reaction patterns. Intuitively, a procedure that aims at reverting time deformation should show its strongest performance with data generated by time deformation. The Monte Carlo simulations show whether this intuition is well grounded.

Denoting economic time by s, we simulate the first-order autoregressive process

$$X_s = \phi X_{s-1} + \varepsilon_s,$$

with Gaussian $N(0,1)$ errors ε. In the basic version of the design, we set $\phi = 0.99$, such that the process becomes strongly correlated but it is still stable.

To this process, we apply one of several time-deformation specifications with different continuous and monotonic one-one functions $g(t)$, $[0,1] \to [0,1]$. The deformation function

$$s = \frac{2}{\pi}\arcsin t, \quad t \in [0,1],$$

accumulates fast at the beginning and end of the year and increases more slowly in 'summer'. The deformation function

$$s = t^{\delta}, \quad t \in [0,1],$$

behaves differently for $\delta < 1$ and $\delta > 1$. For large δ, it speeds up economic activity toward the end of the year, while it focuses on the beginning for $\delta < 1$. We choose $\delta = 0.2$ for a basic design.

In our simulations, we assume that data are generated though not observed at the calendar time represented by t. Observed data are at the 'fine frequency' S_1. The transformation from the basic economic-time process to calendar time works as follows.

As long as $t = g^{-1}(s) < S_1^{-1}$, the observations X_s are aggregated into the first fine seasonal unit, those in the interval $t = g^{-1}(s) \in (S_1^{-1}, 2S_1^{-1}]$ into the second fine unit, and so on.

5.1 Months and quarters

All processes are generated for 5 to 15 years, with some reasonable burn-in. Then, we predict the following year by each of the outlined prediction models: the annual model, the monthly model ($S_1 = 12$), the quarterly model ($S_2 = 4$), and regrouped pseudo-quarters.

Figure 3 gives exemplary time plots of the two basic generating processes. Fine seasonal intervals were set at $S_1 = 12$, while the basic process was generated at 250 units per year. Thus, the design corresponds to business days in economic time and measurement in calendar months. These plots demonstrate the rather extreme patterns of seasonal variation, with activity concentrated in specific parts of the year, while they also emphasize the irregular and non-repetitive nature of the seasonal cycle.

Figure 4 depicts the relative predictive accuracy, with the annual model used as a benchmark. As the sample size increases, simple aggregate annual data (model I) tend to yield acceptable predictions, and gains for any disaggregation become the exception. For small samples, disaggregation effects can be substantial. In the $t^{0.2}$ design, regrouping (model IV) yields similar results as the rival models II and III, while regrouping dominates definitively in the arc-sine design.

Related experiments are depicted in Figure 5, where model III

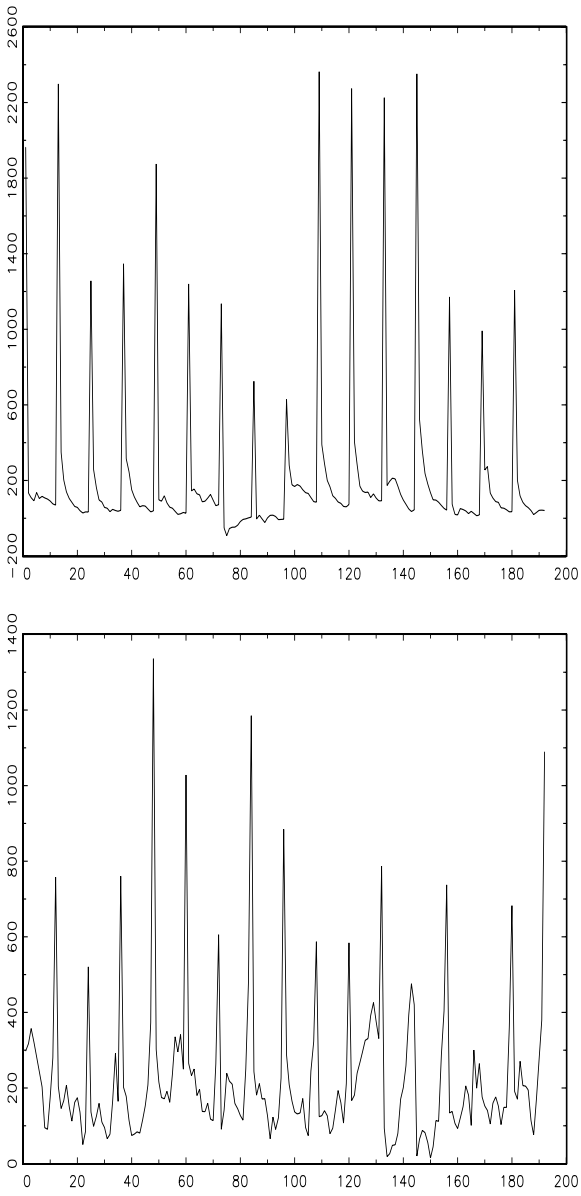

Figure 3: 16 years of generated monthly data. Deformation functions are $s = t^{0.2}, t \in [0,1]$ and $s = \arcsin \frac{2t}{\pi}, t \in [0,1]$.

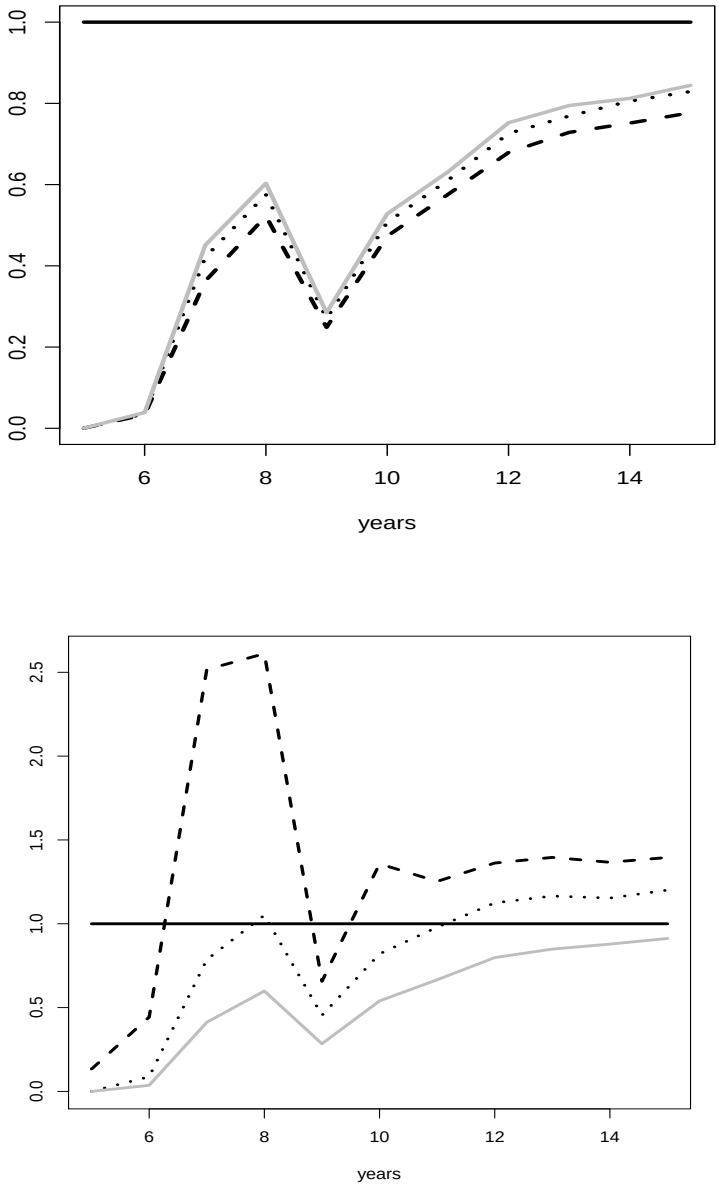

Figure 4: Ratios of prediction MSE for fine seasonal frequency (dashed), crude seasonal frequency (dotted), and regrouping procedure (gray) to the annual model. Generating process is based on $t^{0.2}$ for the top graph and on $\arcsin \frac{2t}{\pi}$ for the bottom graph.

is used as a benchmark. In the left graph, the nonlinearity parameter is allowed to vary from $\delta = 0.2$ to $\delta = 0.6$. Only for the designs with more extreme deformation, gains for regrouping are recognizable. In the right graph, the experiment for the arc-sine generator was repeated with AIC instead of BIC as the lag selection criterion. The regrouping algorithm is substantially better than model III, and even better than the fine-seasons model II, as shown in Figure 4. The AIC version is actually better than the BIC version for most sample sizes, and it also tends to underscore the benefits of regrouping.

In summary, the simulations for the case of months and quarters are a bit disappointing for t^δ deformation and more supportive of regrouping for arc-sine deformation. This distinction matches intuition insofar as the main activity is concentrated toward the end of the year in the latter case, such that the most informative last month typically also becomes the last pseudo-quarter, an important cornerstone for predicting the coming year. Conversely, in the former case the main activity in the first month has much less relevance for the next year.

5.2 Weeks and quarters

All processes are generated for 5 to 15 years, with some reasonable burn-in. Then, we predict the following year by each of the outlined prediction models: the annual model, the weekly model ($S_1 = 52$), the quarterly model ($S_2 = 4$), and regrouped pseudo-quarters. The frequency $S_1 = 52$ was chosen appropriately, such that it is divisible by the quarterly frequency.

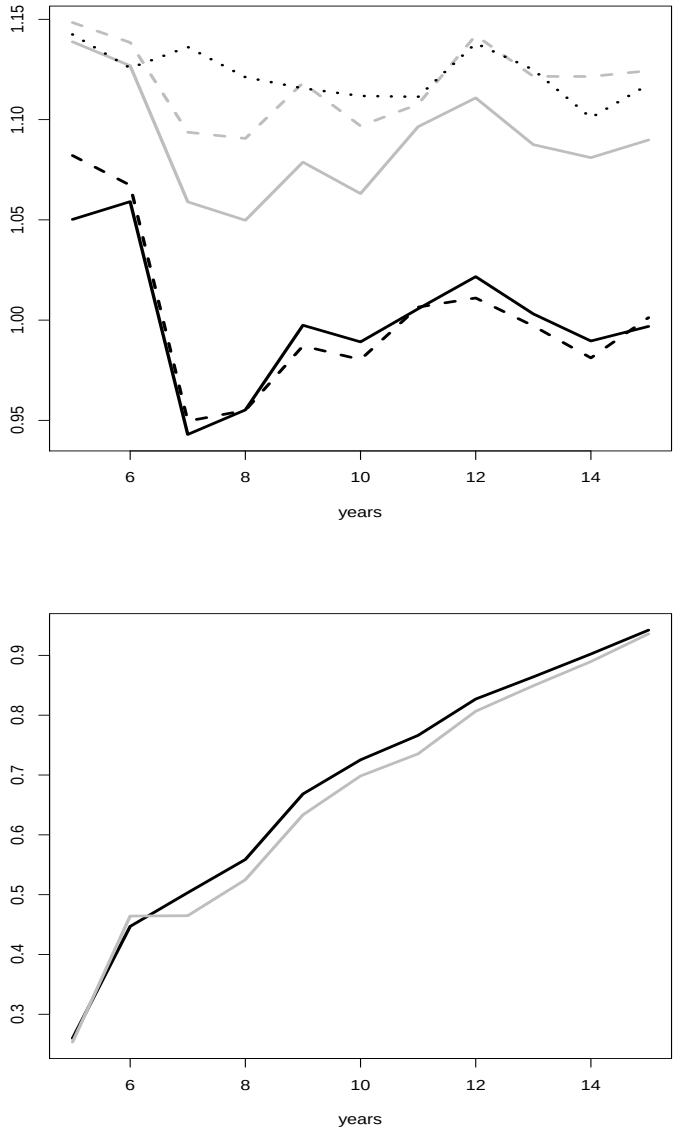

Figure 5: Ratios of prediction MSE for regrouping algorithm to the model using traditional quarters. Upper graph uses $\delta = 0.2$ (black solid), $\delta = 0.3$ (black dashes), $\delta = 0.4$ (gray), $\delta = 0.5$ (gray dashes), and $\delta = 0.6$ (dots), with generating process based on t^{δ}. Lower graph uses arc-sine generating process, with black curve for BIC selection and gray curve for AIC selection.

We note that the basic generating model is identical to the previous subsection but the finest available frequency has changed. Figure 6 shows that the additional information from the weekly observations is difficult to exploit. The weekly model is inferior at all sample sizes. For the arc-sine generating model, this inferiority is so pronounced that its MSE ratio had to be omitted from the graph. By contrast, the regrouping algorithm works well. For geometric deformation, it achieves the values set by the annual model, thus beating the forecast based on calendar quarters by a wide margin. For arc-sine deformation, regrouping is preferable to all rivals, including the annual model.

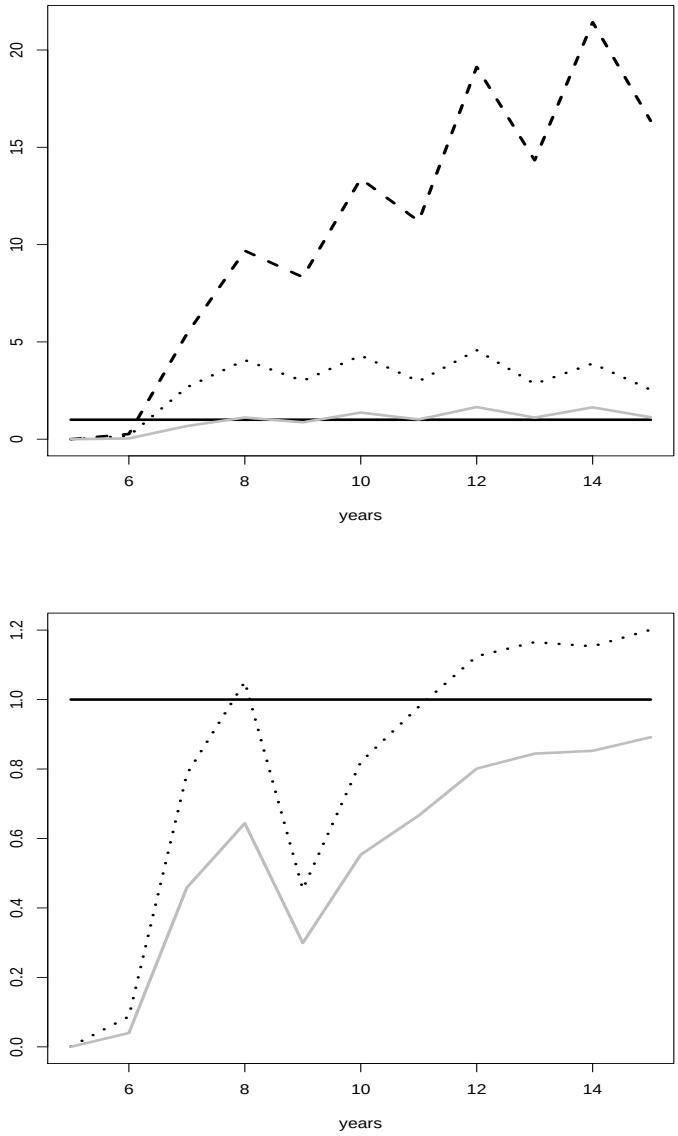

Figure 6: Ratios of prediction MSE for disaggregated prediction models to the annual model. Generating process uses $\delta = 0.2$ in the upper plot, and the arc-sine model in the lower plot. Solid line represents the annual prediction, dashed curve is for forecast based on weeks (upper plot only), dotted curve for quarters, and gray curve for regrouped pseudo-quarters.

6 An empirical application

Weekly data on numbers of people who are under contract of Randstad temporary staffing services are available for the years 1967–2008, i.e. for 42 years. When years had 53 weeks, the data have been adjusted such that 52 weeks per year emerge throughout. Figure 7 displays a time-series plot of the Randstad data.

The trending impression of the data insinuates some transformation, as estimated autoregressions may tend to have unstable roots, which is often detrimental for prediction. For this reason, we also considered first differences of the original data, as they are shown in Figure 8. Clearly, the extremely leptokurtic and heteroskedastic appearance of this data suggests that its prediction could be difficult.

First, we apply the forecasting procedures as outlined above to the original data. We generate out-of-sample one-step autoregressive forecasts for the last 14 years of the sample, using expanding windows. Thus, the forecast for the year 2008 uses 41 years of observations, while the forecast for the year 1995 uses 28 years. All empirical results are summarized in Table 1.

Averaging squared errors across the 14 predicted annual forecasts yields the expected large numbers. In relative terms, the forecast based on weekly observations has an MSE that is 1.9 times as large as the annual forecast MSE, while the forecast based on quarters has an MSE that is only 0.31 the annual forecast MSE. According to this crude evaluation, the regrouping algorithm wins with 0.25 times the annual MSE. However, the variation across years is so sizeable that this summary measure

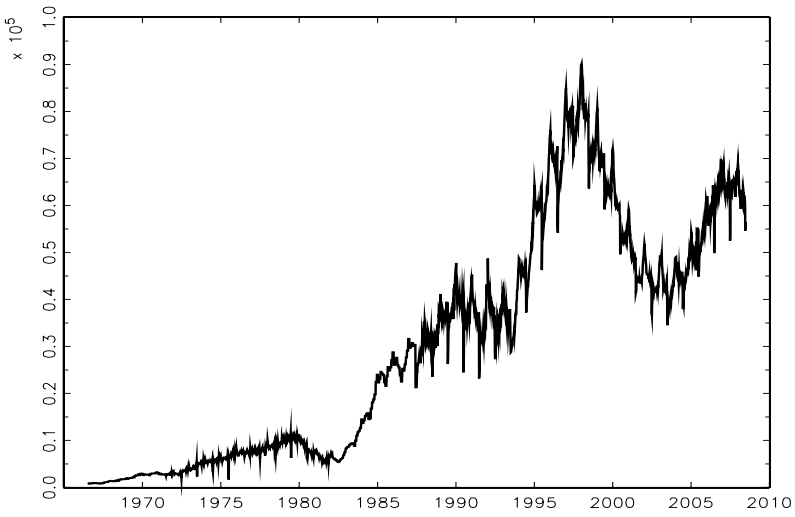

Figure 7: Weekly data for Randstad staffing services for the years 1967–2008.

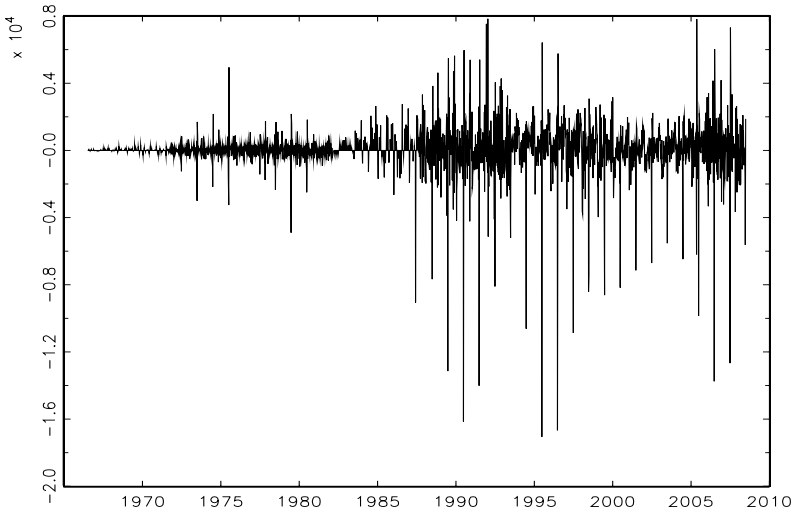

Figure 8: First differences of the Randstad data depicted in Figure 7.

Table 1: Forecasting annual values 1995–2008 of the Randstad data.

	annual	weeks	quarters	pseudo-quarters
levels:				
MSE	8.7e+10	1.6e+11	2.7e+10	2.2e+10
# wins	2	3	4	5
ave. rank	2.86	3.07	2.14	1.93
differences:				
MSE	4.61e+7	4.99e+7	3.77e+7	3.82e+7
# wins	2	5	3	4
ave. rank	2.64	3.50	2.00	1.86

Note: MSE is the average squared error across the predicted years; '# wins' is the number of cases where the respective model achieves the smallest error; 'ave. rank' is the average rank across all 14 cases.

is unreliable. It is of more interest that the regrouping technique achieves a better accuracy than calendar quarters in 9 out of 14 years. The regrouping algorithm loses the duel in the years 2002–2003 and 2005–2007. In 5 years, it achieves the best forecast of all four models, while it never comes in last.

Next, we apply the same prediction models to the differenced data. In a naive average MSE evaluation, calendar quarters and regrouped quarters achieve 0.81 and 0.82 of the annual average MSE. Again, performance across years is quite heterogeneous. In a direct comparison, regrouped predictions are better than calendar quarters in 10 out of 14 cases and are best of all four models in four cases. This time, it is the earlier years that show a slight preference for traditional calendar quarters.

7 Summary and conclusion

We demonstrate that the potential benefits of 'seasonal gerrymandering' in the sense of regrouping higher-frequency observations into lower-frequency aggregates that conform to economic time rather than calendar time depend on the underlying data-generating process. Regular or even deterministic seasonal variation yields poor prospects for such procedures, while existing seasonal time deformation may be more supportive.

In our simulation experiments, we find that seasonal regrouping yields good results for prediction within specific sample-size windows of less than ten years. Whereas the value of this time horizon may be sensitive to the assumed autocorrelation, the pattern is likely to be systematic. In large samples, fine-frequency structures can be estimated consistently and reliably, such that the finest frequency often entails the most precise prediction. In very small samples, simple models with low parameter dimension typically yield the best forecasts. The window between the small-sample and the large-sample case is of interest here.

In our empirical application, we see some benefits for the regrouping procedure, which may indicate that the sample is actually within the mentioned size window. We opine that the many irregularities of the data example are typical for similar applications to economics data.

References

[1] Burns, A.F., and Mitchell, W.C. (1949) *Measuring Business Cycles*. NBER, New York.

[2] Clark, P.K. (1973) 'A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices,' *Econometrica* **41**, 135–155.

[3] Ghysels, E., C. Gourieroux, and J. Jasiak (1995) 'Trading Patterns, Time Deformation and Stochastic Volatility in Foreign Exchange Markets,' Working paper, University of Montreal.

[4] Jiang, H., Gray, H.L., and W.A. Woodward (2006) 'Time-frequency analysis—$G(\lambda)$-stationary processes,' *Computational Statistics & Data Analysis* **51**, 1997–2028.

[5] Jordà, Ò., and M. Marcellino (2004) 'Time-scale transformations of discrete time processes,' *Journal of Time Series Analysis* **25**, 873–894.

[6] Man, K.S. (2004) 'Linear prediction of temporal aggregates under model misspecification,' *International Journal of Forecasting* **20**, 659–670.

[7] Stock, J.H. (1987) 'Measuring Business Cycle Time,' *Journal of Political Economy* **95**, 1240–1261.

[8] Stock, J.H. (1988) 'Estimating Continuous-Time Processes Subject to Time Deformation,' *Journal of the American Statistical Association* **83**, 77–85.

[9] WORKING, H. (1960) 'Note on the Correlation of First Differences of Averages in a Random Chain,' *Econometrica* **28**, 916–918.

Appendix

This appendix contains detailed derivations for the examples of Section 2. Most of them rely on the feature that prediction based on data at the generating frequency entails a straightforward evaluation of variances, while first differences of annual data follow first-order moving-average processes. The minimum forecast variance is then slightly smaller than the variance of the first differences. The role-model case of the Lemma 1 can be used with little variation in all examples.

Lemma 1 *Assume the random walk with independent increments* $X_{\tau,w} = X_{\tau,w-1} + \varepsilon_t$ *and its annual aggregate* $Z_\tau = \sum_{w=1}^{S} X_{\tau,w}$. *The forecast error variance using the annual aggregate is given as*

$$\frac{S(S^2-1)^2\sigma_\varepsilon^2}{6\{2S^2+1-S\sqrt{3(S^2+2)}\}},$$

denoting $\sigma_\varepsilon^2 = \mathrm{var}\varepsilon_t$.

Proof. With an insubstantial modification, this is the situation analyzed by WORKING (1960), who obtained the main result that $Z_\tau - Z_{\tau-1} = \eta_\tau$ follows a first-order moving-average process $\eta_\tau = \xi_\tau + \theta\xi_{\tau-1}$ with first-order correlation

$$\rho_1 = \frac{S^2-1}{2(2S^2+1)}$$

and variance

$$\text{var}(\eta_\tau) = \frac{S(2S^2+1)}{3} = \sigma_\eta^2.$$

For some of our arguments, it is convenient to note that this expression follows from a triangular weighted sum of errors via

$$\sigma_\eta^2 = \sigma_\varepsilon^2(\sum_{w=1}^{S} w^2 + \sum_{w=1}^{S-1} w^2),$$

a two-sided weighted sum of error variances at the generating frequency. Solving the quadratic equation $\rho_1 = \theta/(1+\theta^2)$ yields

$$\theta = \frac{2S^2+1-S\sqrt{3(S^2+2)}}{S^2-1}.$$

The variance $\text{var}(\xi_\tau) = \sigma_\xi^2$ corresponds to the minimum forecast variance and evolves from evaluating

$$\sigma_\xi^2 = \frac{\sigma_\eta^2}{1+\theta^2}.\blacksquare$$

WORKING (1960) remarks that, for larger S, ρ_1 approaches 0.25, and even the smallest $S = 2$ yields $\rho_1 = 1/6$. Also note that $\theta \approx \rho_1$, and that σ_ξ^2 is only slightly smaller than σ_η^2, the forecast variance due to the naive forecast Z_τ that incorrectly assumes that it follows a random walk.

Example 1. First assume disaggregated data are available. Then

$$\begin{aligned}\hat{Z}_{\tau+1} &= \text{E}(Z_{\tau+1}|X_{\tau,s},\ldots,X_{\tau,1},\ldots) \\ &= \text{E}(\sum_{w=1}^{S} X_{\tau+1,w}|X_{\tau,s},\ldots) = SX_{\tau,S},\end{aligned}$$

and thus

$$\begin{aligned} & \mathrm{E}(Z_{\tau+1} - \hat{Z}_{\tau+1})^2 \\ = \; & \mathrm{E}\{SX_{\tau,S} + \sum_{w=1}^{S}(S-w+1)\varepsilon_{\tau+1,w} - SX_{\tau,S}\}^2 \\ = \; & \mathrm{E}\{\sum_{w=1}^{S}(S-w+1)\varepsilon_{\tau+1,w}\}^2 = \sigma_\varepsilon^2 \sum_{w=1}^{S} w^2. \end{aligned}$$

If only annual data are available, Lemma 1 can be applied directly. From the proof of the Lemma, observe that σ_ξ^2 considerably exceeds the above expression that is a one-sided weighted sum. The correction factor is too close to one to compensate this discrepancy.

Example 2. In this case, $Z_\tau - Z_{\tau-1}$ is independent white noise at the annual frequency, and both the forecast for annual and for disaggregated data are clearly given as $\hat{Z}_{\tau+1} = Z_\tau$.

Example 3. Denote the information set formed by past observations $\{X_s, s \leq \tau\}$ by $H_\tau(X)$. By definition,

$$Z_\tau = \sum_{w=1}^{S}(\delta_w + \varepsilon_{\tau,w})$$

and hence

$$\mathrm{E}(Z_{\tau+1}|H_\tau) = \sum_{w=1}^{S}\delta_w.$$

The conditional expectation is identical if data $X_{\tau,w}$ are available, and hence also the concomitant prediction error variance does not change.

Example 4. First consider the annual variable. All calculations closely follow the proof of Lemma 1. Here,

$$Z_{\tau+1} - Z_\tau = \sum_{w=1}^{4}\varepsilon_{\tau+1,w} - \sum_{w=1}^{2}\varepsilon_{\tau,w},$$

an MA(1) process $\eta_\tau = \xi_\tau + \theta\xi_{\tau-1}$ with variance $6\sigma_\varepsilon^2$, first-order covariance $-2\sigma_\varepsilon^2$, and hence first-order correlation $\rho_1 = -1/3$. The implied MA coefficient follows from equating

$$\frac{\theta}{1+\theta^2} = -\frac{1}{3},$$

which yields $\theta = -0.5 * (3 - \sqrt{5}) \approx -0.382$. The MSE is the variance of the implied white noise ξ_t or

$$\sigma_\xi^2 = 6\sigma_\varepsilon^2/(1+\theta^2) \approx 5.236\sigma_\varepsilon^2.$$

The conditional expectation of $Z_{\tau+1}$ based on quarterly data is $X_{\tau,1} + X_{\tau,2}$, which yields a prediction error of just

$$Z_{\tau+1} - X_{\tau,1} + X_{\tau,2} = \sum_{w=1}^{4} \varepsilon_{\tau+1,w},$$

with variance $4\sigma_\varepsilon^2$.

Example 5. The generating process

$$X_{\tau,w} = \begin{cases} X_{\tau-1,2} + \varepsilon_{\tau,1}, & w = 1, \\ X_{\tau,1} + \varepsilon_{\tau,2}, & w = 2, \\ \varepsilon_{\tau,w}, & w = 3, 4, \end{cases}$$

has the forecast based on quarterly data

$$\mathrm{E}(Z_{\tau+1}|X_{\tau,4}, X_{\tau,3}, \ldots) = 2X_{\tau,2}$$

with error variance

$$\begin{aligned} \mathrm{E}(Z_{\tau+1} - 2X_{\tau,2})^2 &= \mathrm{E}(2\varepsilon_{\tau+1,1} + \varepsilon_{\tau+1,2} + \varepsilon_{\tau+1,3} + \varepsilon_{\tau+1,4})^2 \\ &= 7\sigma_\varepsilon^2. \end{aligned}$$

By contrast, if only annual data are available, $Z_\tau - Z_{\tau-1} = \eta_\tau$ again follows a first-order MA process

$$\eta_\tau = 2\varepsilon_{\tau,1} + \sum_{w=2}^{4} \varepsilon_{\tau,w} + \varepsilon_{\tau-1,2} - \sum_{w=3}^{4} \varepsilon_{\tau-1,w},$$

with variance $\sigma_\eta^2 = 10\sigma_\varepsilon^2$, $\rho_1 = -0.1$, and concomitant $\theta \approx -0.1010$. This results in a prediction error variance of around $9.899\sigma_\varepsilon^2$. The two regrouping variants can be evaluated similarly but calculations become a bit involved. As outlined in the text, they were determined by calculation and confirmed by Monte Carlo at $7.84\sigma_\varepsilon^2$ for semesters and at slightly above $9\sigma_\varepsilon^2$ for grouping into the first quarter and the remaining quarters as pseudo-semesters.

Authors: Robert M. Kunst, Philip Hans Franses

Title: Asymmetric Time Aggregation and its Potential Benefits for Forecasting Annual Data

Reihe Ökonomie / Economics Series 252

Editor: Robert M. Kunst (Econometrics)
Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996

Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 59991-555 • http://www.ihs.ac.at

ISSN: 1605-7996