HELLENIC REPUBLIC

UNIVERSITY OF MACEDONIA

Interdepartmental Program of Postgraduate Studies

in Economics

Thesis

THE EFFECT OF AGGREGATION AND THE LOG TRANSFORMATION ON FORECASTING

Koutsigka Myrsini

Supervisor: Pantelidis Theologos

Thessaloniki 2014

Contents

Abstract

Many economic variables are used in logarithms for forecasting and estimation analysis, as this transformation is considered to create a more homogenous variance. The aim of this study is to investigate whether applying logs to the variables of interest improves forecasting precision. Forecasts based on logs are compared to forecasts based on the original series, both for univariate time series and aggregates. Furthermore, forecasts of the aggregates are compared to those obtained by aggregating the forecasts of the individual components based on levels and on logs. The results show that for univariate time series, using logs can be beneficial for forecasting as long as it stabilizes the variance. Otherwise, applying logs can provide inferior forecasts. When it comes to aggregated variables, using logs is often beneficial, especially for long forecast horizons. If logs are not applied, it is preferable to aggregate the forecasts of the components rather than forecasting the aggregate directly only for long forecast horizons.

1. Introduction

In economic studies a usual approach in estimation and forecasting analysis is applying logarithms to the variables of interest. This transformation is often preferred by analysts as it reduces heteroskedasticity effects and stabilizes the variance of a series. If we use for forecasting a model based on the logs of a variable, the forecasts can be transformed through exponentiation to forecasts of the variable in levels. Therefore, it is important for forecasters to know whether applying logs improves forecasting precision, as it is not clear if forecasts based on logs are more accurate than forecasts of a variable in levels. In this study we examine the effect of the logarithmic transformation on the forecasting accuracy of univariate times series, as well as aggregates. A lot of work has been done about aggregated data and the way we can obtain optimal forecasts for the aggregate. However, the majority of those studies refer to linear transformations, and given that the logarithmic transformation is a nonlinear one, the effect of using logs in forecasting aggregated variables has to be investigated.

In this study, we organize two Monte Carlo simulation experiments and we also present the results from two different empirical applications. In the first set of simulations, the effect of the logarithmic transformation on forecasts of a univariate time series is examined by constructing the variable of interest with two different data generation processes (DGPs). In the first one, the log transformation stabilizes the variance of the variable in levels while in the second one the variable in levels already has a stable and homogenous variance. We compute and compare three different forecasts. The first one comes from forecasting directly the variable in levels, the second one by applying logs and then transforming the forecast to obtain the forecast for the original variable and the third one by applying logs and then transforming the forecast using a correction term that theoretically provides a more efficient forecast. The results favor the use of logarithms in forecasting, as long as it stabilizes the variance. If the variance is already stable, applying logs can be harmful for the forecast precision.

In the second simulation experiment, the effect of the logarithmic transformation on aggregates is examined by constructing an aggregated variable that is the sum of two individual components. Four different forecasts of the aggregate are compared: the first one uses directly the aggregated series, the second one applies logs to the aggregate and then transforms the forecast to obtain the forecast of the variable in levels, the third one aggregates the direct forecasts of the components and the last one aggregates the forecasts of the components based on logs. The results show that applying logs can be beneficial for forecasting accuracy and the gains can be substantial for long forecast horizons. If logs are not applied, then aggregating the forecasts of the components can be beneficial only for long forecast horizons.

The study is organized as follows. In section 2 both theoretical and empirical results from the literature about aggregation and the logarithmic transformation are reviewed containing propositions from asymptotic theory about predictors for an aggregate. In section 3 simulation experiments are set up and the results are reported. In section 4 two empirical comparisons concerning real economic data are reported in order to support the simulation results. Section 5 concludes and detailed results tables are given in the Appendix.

2. Literature review

2.1 Literature review on the effect of aggregation on forecasting

Forecasting macroeconomic variables across a large number of countries is a difficult but standard task for economic analysts. Most of those variables are contemporaneous or temporal aggregates, therefore the major problem that arises is the way we can calculate optimal forecasts for the aggregated variables. Different predictors are proposed and compared in the existing literature, however two ways of proceeding dominate. The first one calculates forecasts using directly the aggregated variable, while the second one aggregates the forecasts of the individual components, which are called disaggregates. Forecasts of the disaggregates can be obtained either by univariate predictors for the individual components or by constructing a multivariate model.

We now present the basic theoretical results about contemporaneous aggregation using a notation similar to the one proposed by Lütkepohl (1987). If we consider the variables $x_{1t}, x_{2t}, ... x_{Kt}$ at time t, their contemporaneous aggregate is the sum or weighted sum $y_t = f_1 x_{1t} + f_2 x_{2t} + ... + f_K x_{Kt}$, where $f_1, ... f_K$ are the aggregation weights. Thus, we can write $y_t = F x_t$, where $y_t = (y_1, ... y_M)'$, $x_t = (x_1, ... x_K)'$ and F is an (MxK) transformation matrix of rank M. We consider two different cases, depending on whether the stochastic processes x_t and y_t are completely known or estimated.

The first case investigates possible predictors of y_t under the assumption that x_t and y_t are completely known. Let x_t be a zero mean K – dimensional, stationary, non deterministic stochastic process with moving average (MA) representation

$$x_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} = \Phi(L) u_t, \qquad \Phi_0 = I_K \qquad (1)$$

where u_t is white noise with nonsingular covariance matrix Σ_u and $\det(\Phi(z)) \neq 0$, for $|z| < 1$. If univariate models are used for the components x_{kt} of x_t, then x_{kt} have MA representations

$$x_{kt} = \sum_{i=0}^{\infty} \xi_{ki} w_{k,t-i} = \xi_k(L) w_{kt}, \quad \xi_{k0} = 1, \quad k = 1, ..., K \qquad (2)$$

where $\xi_k(z) \neq 0$ for $|z| < 1$ and w_{kt} are white noise processes. A linear transformation of x_t is $y_t = F x_t$, where F is an (MxK) matrix of rank M, y_t is stationary, nondeterministic with zero mean and MA representation

$$y_t = \sum_{i=0}^{\infty} \Psi_i v_{t-i} = \Psi(L) v_t, \quad \Psi_0 = I_M \qquad (3)$$

and v_t is M – dimensional white noise with covariance matrix Σ_v and $\det(\Psi(z)) \neq 0$ for $|z| < 1$. We consider the following three possible predictors for y_t:

A. An optimal predictor for x_t is obtained based on a multivariate process and the resulting forecasts are transformed.

$$y_t^o(h) = F \sum_{i=0}^{\infty} \Phi_{h+i} u_{t-i} = F x_t(h) \qquad (4)$$

B. An optimal predictor based directly on the aggregate process y_t.

$$y_t(h) = \sum_{i=0}^{\infty} \Psi_{h+i} v_{t-i} \qquad (5)$$

C. Optimal univariate predictors for the individual components of x_t are obtained and then transformed in order to forecast y_t.

$$y_t^u(h) = F x_t^u(h) \qquad (6)$$

where

$$x_t^u(h) = \sum_{i=0}^{\infty} \Xi_{h+i} w_{t-i} \qquad (7)$$

is the predictor for the univariate component.

Wei and Abraham (1981) prove that

- $E\left(y_{t+h} - y_t^o(h)\right)^2 \leq E\left(y_{t+h} - y_t(h)\right)^2$
- $E\left(y_{t+h} - y_t^o(h)\right)^2 \leq E\left(y_{t+h} - y_t^u(h)\right)^2$

That is the optimal predictor is the one based on the joint forecasting model for the entire system, as it uses the largest information set, although the differences between the predictors vanish for long range forecasting (Lütkepohl (1987)). Therefore, aggregating the predictions from the multivariate model is preferable. Their results hold for linear forecasts and do not depend on specific assumptions. However, constructing a multiple time series model may lead to high – dimensional systems and misspecifications (Lütkepohl (1984)). In practice, methods B and C are more common for predicting the aggregate, however it can be shown with some numerical examples that there is no general unconditional inequality for the Mean Squared Errors (MSEs) of the last two predictors. It depends on the data generation process and the transformation matrix (Wei and Abraham (1981), Lütkepohl (1984), Sbrana (2012)). Sbrana and Silvestrini (2009) study a bivariate vector moving average (VMA) process of order 1, comparing the predictors in methods A and B, and show that the forecasting performance of the disaggregate process is superior if the univariate time series are independent. As the components become more dependent, the forecasting ability of the disaggregate process deteriorates. Giacomini and Granger (2002) raise the issue of spatial correlation and show that in such cases aggregating the univariate forecasts leads to higher MSE and thus inaccurate forecasts, due to ignorance of the relationships between regions. Tiao and Guttman (1980) study a stationary multiple MA(q) model and prove that there are no improvements in forecasting the aggregate by its components for h-step ahead forecasts, where h>q. For forecast periods $1 \leq h \leq q$, there are gains in forecasting the components of the aggregate, but the benefits depend on the estimated parameters and the covariance matrix. However, they show that when it comes to non-stationary series, forecasting the components is beneficial for all forecast horizons, but the gains are reduced as the forecast horizon becomes longer.

Kohn (1982) and Lütkepohl (1984) give necessary and sufficient conditions for equality of the MSEs that can be summarized in the following proposition proven by Lütkepohl (Lütkepohl (1987), Proposition 4.1, p. 105):

<u>Proposition 1</u>

Let x_t be a K – dimensional stochastic process as in (1) with univariate subprocesses as in (2). Also let F be an (MxK) matrix of rank M, $y_t = Fx_t$ as in (3) and the predictors $y_t^o(h)$, $y_t(h)$ and $y_t^u(h)$ as in (4), (5) and (6) respectively. Then:

i. $y_t^o(1) = y_t(1) \quad \Leftrightarrow \quad F\Phi(L) = \Psi(L)F$

ii. $y_t^u(1) = y_t(1) \quad \Leftrightarrow \quad F\Xi(L) = \Psi(L)F$

iii. $y_t^o(1) = y_t^u(1) \quad \Leftrightarrow \quad F\Phi(L)^{-1} = F\Xi(L)^{-1}$, provided that Φ(L) and Ξ(L) are invertible.

Lütkepohl (1987) shows that if the conditions of Proposition 1 are satisfied, the equality of the MSEs also holds for the h – step predictors, where h > 1. Sbrana and Silvestrini (2009) show that the condition given in (ii), although sufficient, is not necessary for the equality of the corresponding MSEs. Furthermore, they give sufficient conditions for the equality of the MSEs of the predictors in (ii), that do not satisfy the corresponding condition. Giacomini and Granger (2002) prove that if the data satisfy the condition in (i) and the parameters are estimated, forecasting directly the aggregate may be beneficial.

Lütkepohl (1987) shows that when it comes to nonstationary cases, the results mentioned above still hold, with the exception that the differences in MSEs do not vanish as the forecast horizon increases.

The second case investigates possible predictors of y_t, under the assumption that x_t and y_t are unknown and have to be estimated. Lütkepohl (1987) finds that, provided the constraints

$$F\Phi(L) = \Psi(L)F \quad (8)$$

implied by $y_t^o(1) = y_t(1)$ are valid, it is better to use the disaggregated data for estimating the parameters. If the restrictions (8) are not valid, the predictor $y_t(h)$, based on the univariate aggregated process may be preferable than the predictor $y_t^o(h)$, based on the multivariate process. For the comparison of $y_t^o(h)$ and $y_t^u(h)$ the results are the same. If we ignore the restrictions (8) for the coefficients, $y_t^o(h)$ may be inferior to $y_t^u(h)$.

Lütkepohl (1986) shows that if the DGPs are known, under common assumptions, the forecast MSEs of the disaggregated process are smaller than or equal to the MSEs of the aggregated process, even when temporally aggregated processes are considered. Using asymptotic theory, he proves that even when estimated parameters are used, using disaggregated data is still preferable. Using the disaggregated data is even more important because it generally reduces forecast uncertainty due to estimation.

If we denote the asymptotic MSEs of the three predictors $y_t^o(h), y_t(h), y_t^u(h)$ as $\hat{\sigma}_o^2(h), \hat{\sigma}^2(h), \hat{\sigma}_u^2(h)$ respectively, we can conclude that, if estimated parameters are used, in general $\hat{\sigma}_o^2(h) \le \hat{\sigma}^2(h)$ and $\hat{\sigma}_o^2(h) \le \hat{\sigma}_u^2(h)$ will hold for large sample size T and if possible parameter restrictions are taken into account. However, if restrictions for the coefficients are unknown in the estimation procedure, $y_t^o(h)$ may lose its optimality and may be inferior to the other predictors.

As it becomes clear from the results mentioned above, deciding which method to use for forecasting aggregates is basically an empirical issue. The results differ, depending on the variables, the data generation processes and the restrictions taken in each study.

The majority of the research papers tends to support the aggregation of the components' forecasts, under certain restrictions that differ across studies. Flavin et al. (2009) study growth and inflation for the Euro area and show that pooling forecasts from the disaggregated series and then aggregating them leads to a more consistent forecast than forecasting directly the aggregate, as long as aggregation of the disaggregated forecasts is constructed using Gross Domestic Product (GDP) weighted averages of each country's forecasts. Lütkepohl (2010) studies unemployment rate and inflation and concludes that using the disaggregate forecasts may lead to more efficient forecasts as long as the number of the disaggregated series used is small. Including a larger number of components may lead to efficiency losses due to specification and estimation errors. Marcellino et al. (2001), using Euro-area aggregated variables, show that even if we are interested in measuring an aggregate, pooling country-specific forecasts leads in many cases to more accurate forecasts than using the aggregated data. Bermingham and D'Agostino (2010) investigate the advantages of aggregating forecasts in terms of US and Euro Area inflation. Comparing a number of different models they conclude that, as long as the appropriate model is chosen, aggregating forecasts of the components can improve forecasting accuracy. Hendry and Hubrich (2006) propose, apart from the two approaches mentioned above, an alternative method for forecasting, based on the construction of the aggregate model including disaggregates. They show that theoretically including disaggregates directly in the aggregate model is beneficial for forecasting, but when it comes to empirical work, the benefits of this approach depend on the particular model that is selected. In particular, they study forecasts of US and Euro Area inflation and find that the theoretical results hold only for the US. In a similar work, Hendry and Hubrich (2010) compare three models for forecasting US inflation taking into account mis-specification, estimation uncertainty and parameter changes, and show that including disaggregates directly in the aggregate model is beneficial for forecasting. Espasa and Mayo (2012) investigate forecasts of inflation in 3 different economies, US, UK and Euro Area. They show that disaggregation is beneficial for forecasting the aggregate only when restrictions between the components are considered, such as common trends or serial correlation. Duarte and Rua (2005) study the Portuguese Consumer Price Index (CPI) forecasts using three different levels of disaggregation and show that aggregating the forecasts leads to better accuracy than directly forecasting the aggregate up to a 5 month horizon. They also prove that additional gains can be obtained by increasing the disaggregation level. Brüggemann and Lütkepohl (2013) test whether taking into account stochastically varying aggregation weights improves forecasting accuracy. Through empirical examples of NAFTA real GDP growth rate and European M3 growth rate they conclude that including information of the aggregation weights is beneficial and may improve forecasts. Espasa et al. (2001) study the forecasting of European inflation by countries and by sectors, using Harmonized Index of Consumer Prices (HICP) data, and show that disaggregation by sectors is significant for all horizons greater than one month, as it leads to forecasts with smaller bias and variance, whereas a country disaggregation is beneficial only for the short run (one to three periods). Moser et al. (2004) study forecasts of Austrian inflation and find that aggregating the forecasts of the subindices leads to better accuracy than forecasting directly the HICP. This procedure has the additional advantage of consistency between the forecasts of the subindices and the HICP.

However, in some cases, the results favor the aggregate. For instance, Hubrich (2004) investigates whether aggregating forecasts of subcomponents of HICP improves forecasting accuracy of the aggregated HICP. She shows, using a variety of models and model selection procedures, that no significant benefits occur by aggregating forecasts of disaggregate components of HICP in a 12-month horizon. Furthermore, she shows that when forecasts of the subcomponents are aggregated, the forecast bias increases. Benalal et al. (2004) study forecasts of HICP using data of the 4 largest European countries (France, Germany, Italy and Spain) in order to construct a weighted average for the Euro Area HICP and show that aggregating the forecasts of the sub-components does not necessarily improve the forecast of the aggregated variable, although the differences that occur are generally small. Sbrana and Silvestrini (2009) compare, in their empirical application, forecasts of M1 in Italy and they conclude that the direct forecast of M1 leads to a lower Root Mean Squared Error (RMSE) and thus it is preferable.

2.2 Literature review on the effect of the log transformation on forecasting

A method that is used in some of the studies mentioned above, and in economic studies in general, is using the variables in logarithms. This transformation is often preferred because it limits the effects of heteroskedasticity and it often leads to residuals with more homogenous variance. As a result models fitted in logs may have preferable statistical properties or may be more appropriate for the variables of interest. The question that arises is whether this transformation leads to a more accurate forecast of the variable in levels.

Granger and Newbold (1976) prove theoretically that applying a nonlinear transformation to the optimal forecast of a variable may not result in the optimal forecast of the transformed variable. In the case of logarithms, just exponentiating the forecast of the log to get the forecast of the original variable, which they call the naive forecast, may lead to biased estimates. The optimal forecast occurs by multiplying the naive forecast with a factor containing the forecast error variance of the transformed variable. However, Lütkepohl and Xu (2012) find empirically that there are no substantial gains by using the optimal instead of the naive forecast when it comes to empirical forecast comparison. Further comparison of the naive to the optimal forecast based on the theoretical results of Granger and Newbold (1976) has been done by Arino and Franses (2000) who prove that for long forecast horizons, the optimal forecast performs better than the naive and it should be preferred. A reconsideration of this result comes from Lütkepohl and Bardsen (2011) who claim that despite the theoretical advantages, the naive forecast should be preferred in empirical analysis, because no substantial gains occur from the optimal forecast whereas substantial losses may occur at longer forecast horizons. However, both of them refer only to the forecasts based on logs and none of them compares them to the forecasts of the levels.

Lütkepohl and Xu (2012) investigate whether using the logarithmic transformation is better for forecasts of the variable in levels. They use Monte Carlo simulations and an empirical analysis in univariate models to compare different forecasts of economic variables, like stock indices and consumption. They find that there is no reliable criterion for deciding whether to use logs or levels for forecasting a variable. The choice depends on the homogeneity of the series of interest. The results show that the use of the logs can reduce substantially the forecast MSE as long as the variance of the series becomes more stable. If the logarithmic transformation does not improve the stability of the variance, its utilization may lead to substantial losses in the forecasting accuracy. In that case, forecasts based on the original variable should be preferred. Lütkepohl and Xu (2011) also investigate the logarithmic transformation in forecasting inflation rates. They find out that forecasting the logs of the variables does not necessarily improve forecasting performance. Even in the cases that the logs give lower MSE, the gains are relatively small, so the suggestion is to forecast the original variable rather than the log. Additional work has been done about the appropriateness of the logarithmic transformation in Vector Autoregressive (VAR) models**.** Mayr and Ulbricht (2007) compare, in their empirical study, forecasts of GDP in logs and levels using both the naive and the optimal forecast. They conclude that the logarithmic transformation is at best harmless and has to be questioned.

These results about univariate series can be important for forecasting aggregated series. In the previous subsection the presented results favor aggregating the forecasts of the disaggregated series rather than forecasting the aggregate directly. However, these results refer to linear aggregation while the log transformation is a nonlinear one, thus this case needs to be investigated and cannot be covered by the aggregation theory already presented.

3. Monte Carlo Experiment

Three simulated forecast experiments are carried out to evaluate the forecasting accuracy of different predictors of a variable. In the first two simulation experiments we examine the effect of the log transformation on the forecasting accuracy of a univariate time series. In the third simulation experiment we compare forecasts of aggregated time series in order to examine the effect of aggregation and the log transformation. Different forecasts of the variables of interest are compared in terms of MSE.

3.1 The effect of the logarithmic transformation

The set up that we use in the simulation experiments is similar to that proposed by Lütkepohl and Xu (2012).

3.1.1 Linear DGP of the log series

Let $x_t = \log y_t$ be the natural logarithm of the variable of interest, y_t. We use an autoregressive (AR) process of order 1 for the first differences to generate x_t. The DGP has the form

$$\Delta x_t = c + a\Delta x_{t-1} + e_t, \quad t = 1, 2, ...$$

We use two initial values for x_t, $x_0 = x_{-1} = 0$ and 19 parameter values that range from α = -0.9 to α = 0.9 with step 0.1. The term c is a drift term equal to 0.03 and e_t is a Gaussian white noise, that is $e_t \sim i.i.d.\mathcal{N}\left(0, \sigma_e^2\right)$. We compute y_t as the exponential of the generated series x_t, thus $y_t = \exp x_t$. In that way, the variance of the logs of y_t is homogenous. We consider sample sizes of T = 40, 80, 160 and 320 and we generate additionally 12 post-sample values to evaluate the forecasts. Furthermore, we discard 50 values at the beginning of each sample to avoid start-up effects.

Three different forecasts for y_t are computed and compared in the simulation experiment. The first one is the forecast based directly on the variable y_t. This forecast can be denoted as $y_{t+h|t}^{lin}$, and will be referred in the following as the linear forecast. The second one is obtained through x_t. An h – step forecast for x_t is computed and then transformed through exponentiation to the h – step forecast for y_t, which can be denoted as $y_{t+h|t}^{nai} = \exp x_{t+h|t}$. This forecast is referred to as naive by

Lütkepohl and Xu (2012), as it was proven earlier by Granger and Newbold (1976), that this is not the optimal predictor. According to Granger and Newbold (1976), this is a biased forecast, and the optimal one should be $y_{t+h|t}^{opt} = \exp x_{t+h|t} \cdot \exp\left(\frac{1}{2}\sigma_x^2(h)\right)$, where $\sigma_x^2(h)$ is the forecast error variance of x_t. Therefore, the third forecast is the optimal predictor proposed by Granger and Newbold (1976).

Only AR(1) processes with an intercept are fitted[1], and the forecasts are computed for forecast horizons h up to 12 periods. We carry out 5000 replications of the experiment and we compute forecast MSEs. In order to compare the efficiency of the forecasts, we compute ratios of MSEs of naive forecast relative to linear forecast. Ratios smaller than 1 indicate that the MSE of the naive forecast is smaller than the MSE of the linear one, thus the naive forecast is preferable, whereas ratios that exceed unity indicate that the MSE of the naive forecast is larger than the MSE of the linear one, thus linear forecasts should be preferred. Some basic results are presented in Tables 1 to 4. [2]

Table 1: Ratios of forecast MSEs of naive relative to linear forecast for T = 40

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	0.957	0.935	0.873	0.853	1.022	0.640	0.607	0.612	0.741	0.933
3	0.892	0.844	0.749	0.642	1.217	0.448	0.316	0.290	0.471	0.846
6	0.836	0.715	0.600	0.412	0.960	0.300	0.178	0.150	0.276	0.782
12	0.703	0.566	0.443	0.111	0.661	0.154	0.097	0.078	0.135	0.688

Table 2: Ratios of forecast MSEs of naive relative to linear forecast for T = 80

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	0.853	0.820	0.797	0.781	1.263	0.384	0.441	0.547	0.722	0.762
3	0.756	0.638	0.561	0.574	0.845	0.228	0.172	0.266	0.457	0.716
6	0.612	0.477	0.410	0.388	0.517	0.124	0.085	0.117	0.272	0.930
12	0.429	0.315	0.264	0.161	0.306	0.064	0.045	0.050	0.141	1.093

[1] In this way, there is no need to use model selection that can influence our findings.
[2] Full tables and results are reported in the Appendix.

Table 3: Ratios of forecast MSEs of naive relative to linear forecast for T = 160

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	0.739	0.730	0.765	0.883	0.946	0.262	0.374	0.526	0.719	0.914
3	0.583	0.484	0.496	0.757	0.655	0.134	0.130	0.254	0.454	0.892
6	0.417	0.309	0.318	0.471	1.389	0.067	0.055	0.120	0.279	0.779
12	0.255	0.198	0.195	0.249	1.212	0.033	0.027	0.049	0.142	0.664

Table 4: Ratios of forecast MSEs of naive relative to linear forecast for T = 320

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	0.645	0.635	0.723	0.583	0.587	0.196	0.340	0.523	0.697	0.686
3	0.462	0.380	0.433	0.467	0.571	0.093	0.120	0.252	0.440	0.453
6	0.303	0.238	0.286	0.266	0.578	0.043	0.049	0.119	0.272	0.417
12	0.181	0.152	0.169	0.081	0.486	0.022	0.021	0.047	0.134	0.334

For each sample size two different cases are considered, based on two different values of σ_e^2, $\sigma_e^2 = 0.001$ or $\sigma_e^2 = 0.0001$. As it is clear from the results in the tables, almost all ratios are smaller than 1 with few exceptions, for example for T = 40, $\sigma_e^2 = 0.001$, α = 0.9 and forecast horizon h = 1 and h = 3, where the ratios are 1.022 and 1.217 respectively. Thus, the logarithmic transformation is beneficial, and in some cases the gains are substantial. For instance, for T = 320, $\sigma_e^2 = 0.0001$, α = - 0.5 and forecast horizon h = 12, the ratio of the MSE of the naive forecast relative to linear is 0.021, in other words the naive forecast is more than 40 times smaller than the linear one. Furthermore, it can be seen that for the same sample size, the gains from using logs increase for longer forecast horizon, and are even more substantial when σ_e^2 becomes smaller. For instance, for sample size T = 40, parameter value α = - 0.5 and $\sigma_e^2 = 0.001$ the ratio of the MSEs for h = 1 is equal to 0.935, whereas for h = 12 decreases to 0.566, a difference that is quite significant. For the same sample size and parameter value, ratios are even smaller for $\sigma_e^2 = 0.0001$. For h = 1 the ratio of the MSEs is equal to 0.607 and for h = 12 it reaches 0.097. In other words, MSE of the naive forecast is more than 10 times smaller than the linear one. A few exceptions can be noticed in some cases with parameter value α = 0.9, as in Table 2, where the naive forecast for

$\sigma_e^2 = 0.0001$ loses its optimality as the forecast horizon increases and for h = 12 is inferior to the linear one, and in Table 3, for sample size T = 160 and $\sigma_e^2 = 0.001$ where it does not have a stable behavior.

So far the naive forecast is compared to the linear one. We now compare the naive forecast to the optimal one, as proposed by Granger and Newbold (1976). The correction term $\sigma_x^2(h)$ is computed using the following methodology proposed by Lütkepohl and Xu (2012): Let $\hat{\rho}(L)$ be the AR(1) estimated polynomial of the series in first differences, where L is the lag operator. Also assume an AR(2) model for the original series with estimated coefficients $\hat{a}_1$ and $\hat{a}_2$. Values of the coefficients can be calculated from the equation $1-\hat{a}_1 L-\hat{a}_2 L^2 = \hat{\rho}(L)(1-L)$. Then the forecast error variance can be computed as

$$\hat{\sigma}_x^2(h) = \hat{\sigma}_\varepsilon^2 \sum_{i=0}^{h-1} \hat{\varphi}_i^2$$

where $\hat{\varphi}_i = \sum_{j=1}^{\min(i,p+1)} \hat{\varphi}_{i-j}\hat{a}_j$, with initial value $\hat{\varphi}_0 = 1$, $\hat{a}_1$ and $\hat{a}_2$ the estimated coefficients and $\hat{\sigma}_\varepsilon^2 = \frac{1}{\mathrm{T}-1-p}\sum_{t=2}^{T} \hat{\varepsilon}_t^2$ the estimated residual variance.

Ratios of MSEs of optimal forecast relative to naive forecast are presented in Tables 5 to 8. As it can be seen, ratios are close to unity in all sample sizes and for all forecast horizons. However, ratios greater than unity dominate. For instance, for sample size T = 160 , $\sigma_e^2 = 0.001$, parameter value α = 0.9 and 12 – step ahead forecast horizon, the ratio of the MSEs of the optimal relative to naive is 1.600, a number substantially greater than one.

Table 5: Ratios of forecast MSEs of optimal relative to naive forecast for T = 40

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.001	1.001	1.001	1.004	1.009	1.000	1.000	1.000	1.000	1.002
3	1.001	1.002	1.005	1.014	1.058	1.000	1.000	1.001	1.001	1.008
6	1.003	1.004	1.009	1.034	1.126	1.001	1.000	1.001	1.003	1.019
12	1.006	1.009	1.018	1.075	1.364	1.001	1.001	1.003	1.006	1.051

Table 6: Ratios of forecast MSEs of optimal relative to naive forecast for T = 80

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.001	1.001	1.002	1.004	1.028	1.000	1.000	1.000	1.001	1.001
3	1.002	1.003	1.006	1.017	1.070	1.000	1.000	1.001	1.002	1.016
6	1.003	1.005	1.013	1.039	1.161	1.000	1.001	1.001	1.004	1.045
12	1.007	1.008	1.021	1.075	1.452	1.000	1.001	1.002	1.008	1.088

Table 7: Ratios of forecast MSEs of optimal relative to naive forecast for T = 160

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.001	1.002	1.003	1.007	1.034	1.000	1.000	1.000	1.000	1.002
3	1.002	1.003	1.007	1.021	1.078	1.000	1.000	1.000	1.002	1.001
6	1.002	1.006	1.010	1.035	1.200	1.000	1.000	1.001	1.003	1.012
12	1.006	1.010	1.018	1.063	1.600	1.000	1.001	1.001	1.009	1.061

Table 8: Ratios of forecast MSEs of optimal relative to naive forecast for T = 320

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.001	1.000	1.003	1.013	0.982	1.000	1.000	1.000	1.000	0.997
3	1.002	1.002	1.005	1.029	0.931	1.000	1.001	1.000	1.001	1.004
6	1.004	1.005	1.012	1.052	0.924	1.000	1.001	1.000	1.003	1.030
12	1.007	1.011	1.023	1.105	1.212	1.000	1.001	1.001	1.007	1.112

The situation gets worse as the forecast horizon increases, for all sample sizes and parameter values. As a result, the theoretically optimal predictor appears inferior to the naive one. An explanation for this outcome is proposed by Lütkepohl and Xu (2012), who claim that the optimal forecast may be inferior because the forecast error variance that is used for the correction of the bias is unknown and estimated. So when it comes to applied forecasting, using the optimal forecast does not improve the forecasting accuracy and the naive one should be preferred.

To conclude, using the logarithmic transformation can be beneficial for forecasting economic variables, if this transformation leads to a more homogenous variance for the variable of interest.

The optimal predictor proposed by Granger and Newbold (1976) may not be of great value in applied work, so the naive predictor should be preferred.

3.1.2 Linear DGP of the original series

Lütkepohl and Xu (2012) mention that using logs in the variable of interest may induce heteroscedasticity. So the question they bring forward is whether the log transformation can be beneficial for forecasting, if the original variable is already stable. In order to investigate the logarithmic effect when the variance is already stable, the variable of interest y_t is constructed by an AR(1) process for the first differences Δy_t, that is

$$\Delta y_t = c + a\Delta y_{t-1} + e_t, \quad t = 1, 2, ...$$

The initial values that we use for y_t are $y_0 = y_{-1} = 0$ and the drift term is again $c = 0.03$. The error term e_t is independent, identically normally distributed with zero mean and variance σ_e^2. Then series x_t is obtained by applying logs in y_t, that is $x_t = \log y_t$. Samples of size 40, 80 160 and 320 are used, as in the previous simulation, and the forecast horizon h is up to 12 periods. Again, ratios of MSEs of naive relative to linear forecast are computed for the comparison of the two predictors. The results are presented in Tables 9 to 12.

As it can be seen, for all sample sizes and for all forecast horizons, all ratios are greater than 1, and some substantially so. For instance, for sample size T = 80, parameter value α = -0.5, h = 12 and $\sigma_e^2 = 0.0001$, the ratio of the MSE of the naive forecast relative to the linear is 40.447, that is the MSE of the naive is more than 40 times larger than the MSE of the linear forecast. In such a case, using the log transformation can harm the forecast precision, and in some cases this effect can be dramatic. The situation gets worse as the forecast horizon increases and the variance becomes smaller. For example, for sample size T = 40, parameter value α = 0.5 and $\sigma_e^2 = 0.001$, the ratio of the MSEs is 1.140 for forecast horizon h = 1 and reaches 2.997 for h = 12. For the same sample and parameter value, ratios are even larger if the variance is set equal to $\sigma_e^2 = 0.0001$. For the 1 – step ahead forecast the ratio is 1.427 and for the 12 – step ahead forecast it is equal to 11.392, almost 4 times larger than the ratio of the 12 – step ahead forecast for $\sigma_e^2 = 0.001$.

Table 9: Ratios of forecast MSEs of naive relative to linear forecast for T = 40

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.135	1.127	1.124	1.140	1.145	1.882	1.782	1.666	1.427	1.124
3	1.296	1.392	1.385	1.404	1.375	3.093	3.618	3.605	2.407	1.360
6	1.675	1.908	1.881	1.863	1.727	5.854	7.118	6.945	4.551	1.854
12	2.762	3.205	3.138	2.997	2.517	14.153	16.385	15.854	11.392	3.262

Table 10: Ratios of forecast MSEs of naive relative to linear forecast for T = 80

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.315	1.294	1.265	1.269	1.164	3.421	2.899	2.204	1.529	1.093
3	1.734	1.891	1.934	1.768	1.418	6.796	8.273	6.262	2.976	1.273
6	2.500	2.982	3.058	2.684	1.802	14.336	17.859	14.746	6.580	1.617
12	4.884	5.705	5.635	5.107	2.744	33.765	40.447	36.573	18.938	2.545

Table 11: Ratios of forecast MSEs of naive relative to linear forecast for T = 160

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	1.716	1.680	1.605	1.434	1.129	6.762	4.533	2.630	1.502	1.114
3	2.603	3.101	3.078	2.392	1.338	14.517	17.181	9.644	3.019	1.354
6	4.212	5.487	5.358	4.206	1.714	31.964	40.125	26.730	7.044	1.763
12	9.177	11.001	10.703	9.177	2.718	76.535	91.246	70.478	23.737	2.890

Table 12: Ratios of forecast MSEs of naive relative to linear forecast for T = 320

	$\sigma_e^2 = 0.001$					$\sigma_e^2 = 0.0001$				
	AR parameter					AR parameter				
h	-0.9	-0.5	0	0.5	0.9	-0.9	-0.5	0	0.5	0.9
1	2.583	2.355	2.081	1.615	1.148	12.500	6.279	2.809	1.493	1.159
3	4.487	5.442	4.901	3.216	1.370	29.933	31.147	12.426	2.925	1.469
6	8.413	10.699	9.355	6.743	1.776	69.884	74.654	40.881	7.131	2.043
12	17.625	21.362	20.203	15.959	2.833	157.303	178.487	124.387	26.060	3.766

As a result, if the original variable is already stable and the logs do not improve homogeneity in the variance, using the logarithmic transformation can be harmful for forecasting accuracy.

The general conclusion is that the log transformation can improve forecasting accuracy only if it stabilizes the variance of the original variable. If the variance is already stable, the most accurate forecast can be computed by forecasting directly the variable in levels. Furthermore, there are no substantial gains by using the optimal instead of the naive forecast, therefore, the naive should be preferred.

3.2 Logarithmic effect on aggregation

The following simulation experiment concerns variables which are aggregates of a number of component variables. Let y_t be the aggregate of two individual components, which are called disaggregates, y_{1t} and y_{2t}. We assume that $y_t = y_{1t} + y_{2t}$. The main question that rises is in which way one can obtain the most accurate forecast for the aggregate. Forecasts y_{t+h} of the aggregate, for forecast horizon h, can be obtained as follows[3]:

A. A DGP process for the aggregate y_t is constructed and then forecasts $y_{t+h|t}^{lin}$ are computed directly from the aggregated series, which we will henceforth call linear.

B. A DGP process for x_t is constructed, where $x_t = \log y_t$ is the logarithmic transformation of the aggregate y_t. Forecasts $x_{t+h|t}$ are computed for x_t and then forecasts for the aggregate y_t are obtained by exponentiating the forecasts based on logs. Therefore, the aggregate forecasts are $y_{t+h|t}^{nai} = \exp\left(x_{t+h|t}\right)$, which we will henceforth call naive.

C. A DGP is constructed for each one of the disaggregate components y_{1t}, y_{2t}. Forecasts of the aggregate y_t are obtained by simply adding the forecasts of the individual components. Therefore the aggregate forecast is $y_{t+h|t}^{dis} = y_{1,t+h|t} + y_{2,t+h|t}$

D. A DGP process is constructed for x_{1t}, x_{2t}, where x_{1t}, x_{2t} are the logarithmic transformations of y_{1t}, y_{2t}, that is $x_{1t} = \log y_{1t}$ and $x_{2t} = \log y_{2t}$. Forecasts of the aggregate are obtained by adding the forecasts of the components based on logs. That is $y_{t+h|t}^{naidis} = y_{1,t+h|t}^{nai} + y_{2,t+h|t}^{nai}$ where

[3] One could also think of various alternative predictors not considered in our study.

$y_{1,t+h|t}^{nai} = \exp\left(x_{1,t+h|t}\right)$ and $y_{2,t+h|t}^{nai} = \exp\left(x_{2,t+h|t}\right)$ are the forecasts of the components, computed by exponentiating the forecasts of the respective logarithms.

The first model serves as the benchmark. Each of the following models are compared to the first one. We compare the second model with the benchmark in order to examine the effect of the logarithmic transformation on forecasting directly the aggregate. We compare the third model with the benchmark in order to examine the effect of aggregation, in other words if aggregating the forecasts of the individual components can be beneficial for the forecasting accuracy of the aggregate. Finally we compare the last model with the benchmark, in order to examine the combined effect of aggregation and the logarithmic transformation.

The DGP process for the first differences of the disaggregated series in logs is a VAR(1) with parameter values estimated from the actual economic variables in section 4, that is

$$\Delta x_t = \begin{pmatrix} 0.009 \\ 0.007 \end{pmatrix} + \begin{pmatrix} -0.014 & -0.162 \\ 0.005 & 0.102 \end{pmatrix} \Delta x_{t-1} + u_t$$

where $u_t \sim N\left(0, \Sigma_u\right)$, covariance matrix $\Sigma_u = \begin{pmatrix} 0.007 & -9.34\times10^{-5} \\ -9.34\times10^{-5} & 4.35\times10^{-5} \end{pmatrix}$ and initial values $x_1\left(1\right) = 8.1936$, $x_1\left(2\right) = 8.1321$, $x_2\left(1\right) = 11.934$ and $x_2\left(2\right) = 11.9394$, which are taken from the initial values of the data series used in section 4. The disaggregated series are obtained by exponentiating the series constructed by the VAR(1) and the aggregate is obtained by adding the two disaggregate series. In other words, $y_t = y_{1t} + y_{2t}$, where $y_{1t} = \exp x_{1t}$ and $y_{2t} = \exp x_{2t}$. We use sample sizes of T = 40, 80, 160 and 320 and we fit AR(p) processes to the first differences of all variables. The Schwarz (SIC) criterion is used to choose the appropriate lag orders, with maximum lag order equal to 4 in the selection procedure. We carry out 5000 replications of the simulation experiment with forecast horizons h up to 12 periods and we compute the four different forecasts of the aggregate as well as their forecast MSEs that will be used for comparison. For the forecasts obtained through the logarithmic transformation, only the naive forecast is used, as it was shown in the previous section that the optimal forecast does not improve the forecasting accuracy.

In order to compare the forecast of the direct aggregate with the one obtained based on its logarithm, their forecast MSEs are compared by calculating ratios of $MSE\left(y_{t+h|t}^{nai}\right)$ relative to $MSE\left(y_{t+h|t}^{lin}\right)$. The results are reported in Table 13.

It is clear that all ratios are smaller than 1, that is $MSE\left(y_{t+h|t}^{nai}\right) < MSE\left(y_{t+h|t}^{lin}\right)$. Therefore, using the logarithmic transformation to forecast the aggregate is preferable than forecasting the aggregate directly. In many cases the benefits are substantial, as for sample size T = 80, where the 12 – step ahead forecast ratio is 0.550, that is the MSE of the linear forecast is nearly 2 times larger than that of the naive one. Furthermore, the benefits from the logarithmic transformation increase for longer forecast horizons. For instance, for sample size T = 160, the ratio of the 1 – step ahead forecast MSEs is 0.864 and reaches 0.354 for the 12 – step ahead forecasts. This tendency is repeated for all sample sizes and all forecast horizons. These results come to an agreement with the findings presented in the previous section.

Table 13: Ratios of forecast MSEs of the naive relative to linear forecast of the aggregate

Sample size T	Forecast horizon h			
	1	3	6	12
40	0.990	0.958	0.913	0.840
80	0.912	0.809	0.696	0.551
160	0.864	0.701	0.545	0.354
320	0.844	0.884	0.687	0.443

The next comparison between $y_{t+h|t}^{dis}$ and $y_{t+h|t}^{lin}$ examines the question whether it is better to forecast directly the aggregate or aggregate the forecasts of the components. Again we compute ratios of MSEs, in particular ratios of $MSE\left(y_{t+h|t}^{dis}\right)$ relative to $MSE\left(y_{t+h|t}^{lin}\right)$ which are reported in Table 14. Once again ratios are close to unity.

Table 14: Ratios of forecast MSEs of the disaggregate relative to the linear forecast

Sample size T	Forecast horizon			
	1	3	6	12
40	1.015	1.002	1.011	1.006
80	1.018	1.035	1.063	1.115
160	0.997	0.999	0.993	0.999
320	1.002	0.971	0.963	0.950

For samples with size T = 40 or 80, all ratios are greater than one, that is $MSE\left(y_{t+h|t}^{dis}\right) > MSE\left(y_{t+h|t}^{lin}\right)$. So aggregating the forecasts of the components can be harmful for the forecasting accuracy, as it leads to higher forecast MSEs, although the differences are not extremely large, as the ratios are very close to unity. However for larger samples (T = 160 or 320), almost all ratios are smaller than 1, that is $MSE\left(y_{t+h|t}^{dis}\right) < MSE\left(y_{t+h|t}^{lin}\right)$. Thus, forecasting the disaggregate components and then aggregating the forecasts can be beneficial for larger samples.

The last comparison is between $y_{t+h|t}^{naidis}$ and $y_{t+h|t}^{lin}$, so we test whether using logs to forecast the components and then aggregating them can lead to higher forecast accuracy than forecasting the aggregate directly. We compute ratios of $MSE\left(y_{t+h|t}^{naidis}\right)$ relative to $MSE\left(y_{t+h|t}^{lin}\right)$ and we report them in Table 15.

Table 15: Ratios of forecast MSEs of the disaggregate forecast based on logs relative to the linear forecast

Sample size T	Forecast horizon			
	1	3	6	12
40	1.009	0.971	0.939	0.884
80	0.931	0.833	0.728	0.599
160	0.871	0.704	0.545	0.361
320	0.820	0.953	0.771	0.504

Almost all ratios are smaller than 1, so applying logs on the disaggregate series and then aggregating the forecasts is clearly beneficial. An exception appears for sample size T = 40 and forecast horizon h = 1, with the ratio being equal to 1.009, but it is so close to unity that the two forecasts can be considered equivalent. For the same sample and longer forecast horizons, ratios are smaller than 1, that is $MSE\left(y_{t+h|t}^{naidis}\right) < MSE\left(y_{t+h|t}^{lin}\right)$. Thus, using logarithms for forecasting the disaggregate components and then aggregating the transformed forecasts can be beneficial for forecasting the aggregate. Furthermore, the gains are even more substantial for long forecast horizons. For instance, for sample size T = 80, the ratio for 1 – step ahead forecasts is equal to 0.931 and for 12 – step ahead forecast is equal to 0.599. In other words, the difference between $MSE\left(y_{t+h|t}^{naidis}\right)$ and $MSE\left(y_{t+h|t}^{lin}\right)$ increases with h.

From the aforementioned results, it is obvious that using the logarithmic transformation can be beneficial for forecasting the aggregated variable. A question that occurs is whether logarithms may lead to more accurate forecasts when they are applied in the aggregated series, or it is better to apply logs in the disaggregate components and then aggregate the forecasts. In other words, $y_{t+h|t}^{nai}$ and $y_{t+h|t}^{naidis}$ should be compared. Following the approach used in the above comparisons, ratios of $MSE\left(y_{t+h|t}^{naidis}\right)$ relative to $MSE\left(y_{t+h|t}^{nai}\right)$ are computed and presented in Table 16.

Table 16: Ratios of forecast MSEs of disaggregate forecast based on logs relative to naive forecast

Sample size	Forecast horizon			
T	1	3	6	12
40	1.019	1.014	1.029	1.053
80	1.020	1.030	1.046	1.087
160	1.008	1.004	1.000	1.020
320	0.972	1.078	1.124	1.140

Almost all ratios are greater than one, that is $MSE\left(y_{t+h|t}^{naidis}\right) > MSE\left(y_{t+h|t}^{nai}\right)$. So the logarithmic transformation can be more beneficial if it is applied to the aggregate series y_t rather than to its components. The improvements in forecasting accuracy increase for longer forecast horizons. For example, for sample size T = 40, the ratio for 1 – step and 12 – step ahead forecasts is 1.019 and 1.053 respectively.

In summary, it seems that when it comes to forecasting an aggregated series, using the logarithmic transformation either on the aggregate or the disaggregate series improves forecasting accuracy in terms of MSE. Nevertheless it is better to apply the logs in the aggregated series and not in the components. If the logarithmic transformation is not applied, then forecasting the disaggregate components and then aggregating the forecasts can be beneficial for large samples .

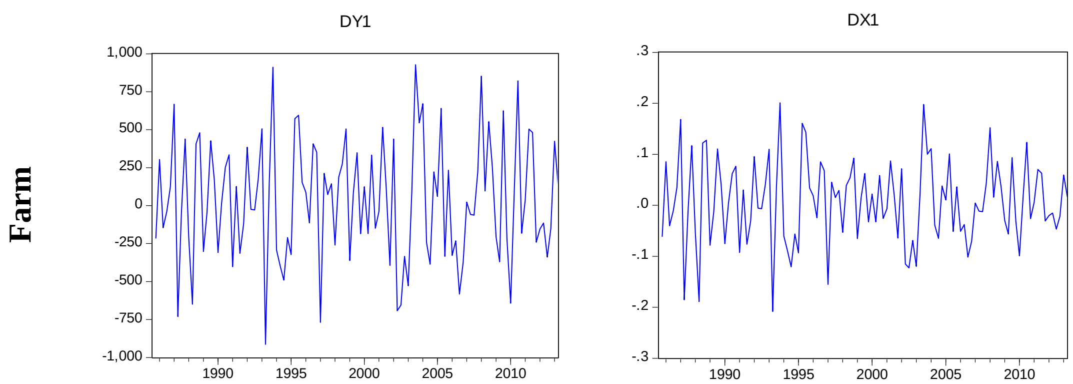
4. Empirical Application

In this section two empirical applications are carried out in order to discuss the results of applying the 4 different approaches on forecasting aggregates that are presented in the previous section. The variable of interest in both empirical applications is real GDP, as logs of GDP are often considered in economic analyses. In both applications we use the rolling scheme for estimation and forecasting and we compare the four different forecasts of the aggregates in terms of MSE.

4.1 Australian GDP

Quarterly data on real GDP measured in millions of Australian dollars are considered, starting in 1985Q3 until 2013Q2. The disaggregate components are farm and non – farm GDP, where non – farm refers to GDP less gross farm output. The data source is the Reserve Bank of Australia and the aggregated Australian real GDP is computed by aggregating farm and non – farm real GDP. The first differences and first differences of logs of the aggregate and its components are plotted in Figure 1. We can observe that applying logs does not improve stability in the variation of the variables of interest.

Augmented Dickey Fuller (ADF) tests are carried out for full samples that do not reject the presence of a unit root. However non – stationarity is rejected in the first differences of all variables, so they are assumed to be I(1). Thus, AR(p) processes are fitted to the first differences, where the order p is chosen by the SIC criterion with maximum lag order equal to 4. We compute four different forecasts of the aggregate proposed in the third simulation experiment and we consider forecast horizons up to 12 periods.



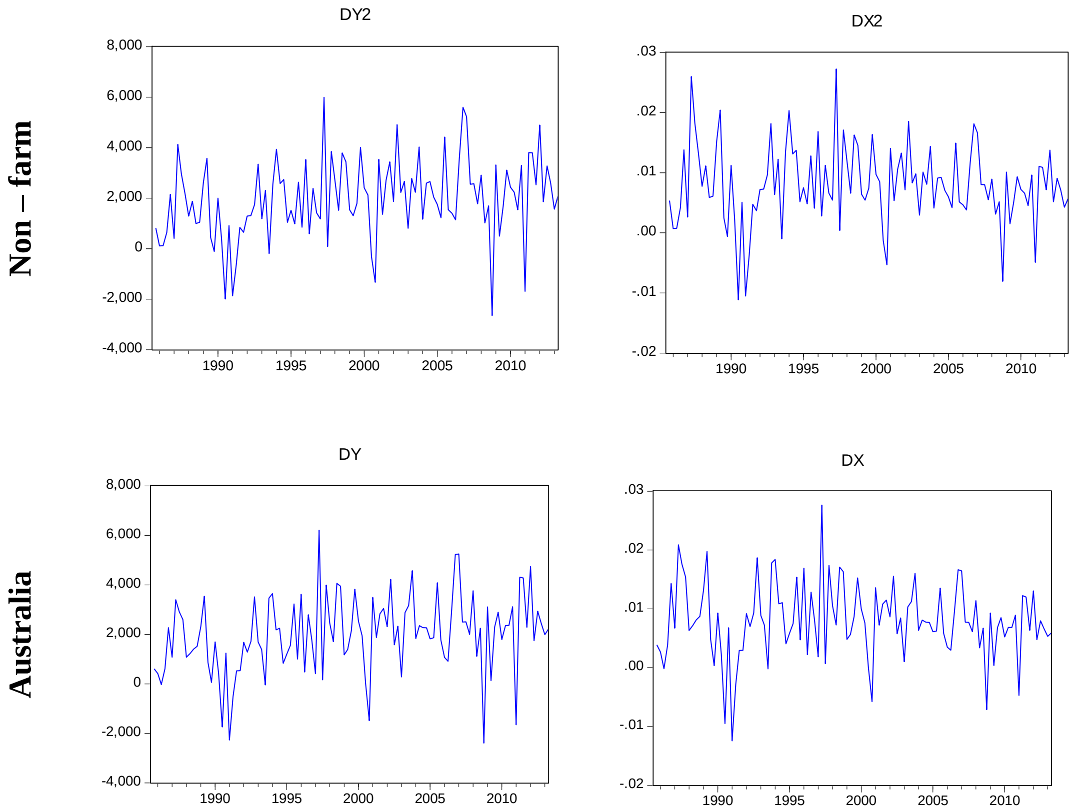


Figure 1: First differences and first differences of logs of real GDP, where Y1 is the farm real GDP, Y2 is the non - farm real GDP, Y is the aggregate real GDP and X1, X2, X the respective logs

Table 17 reports the ratios of forecast MSEs of the forecast obtained by applying logs in the aggregate relative to the forecast obtained directly by the aggregate. Clearly all ratios are greater than one. Therefore, applying logs to the aggregated series in order to obtain a forecast can be harmful for the forecast precision. This holds for all forecast horizons. Furthermore it is noticeable that this loss of forecast precision becomes more substantial as the forecast horizon increases. As it can be seen, for 1 – step ahead forecasts the ratio is 1.067 whereas for 12 – step ahead forecasts it rises to 1.573. As a result, forecasting directly the aggregate can be beneficial, in terms of MSE.

Table 17: Ratios of forecast MSEs of naive relative to linear forecast

Forecast horizon	$\left(\frac{MSE\left(y^{nai}_{t+h\|t}\right)}{MSE\left(y^{lin}_{t+h\|t}\right)}\right)$
1	1.067
3	1.108
6	1.265
12	1.573

We continue with the comparison of the forecast obtained by aggregating the forecasts of the components with the one obtained by the aggregate directly. Ratios of the corresponding MSEs are presented in Table 18.

Table 18: Ratios of forecast MSEs of disaggregate relative to linear forecast

Forecast horizon	$\left(\frac{MSE\left(y_{t+h\mid t}^{dis}\right)}{MSE\left(y_{t+h\mid t}^{lin}\right)}\right)$
1	1.023
3	0.986
6	0.992
12	1.007

In this case the results are ambiguous. For h = 1, ratio is 1.023, that is forecasting the aggregate directly can lead to lower MSE. For h = 3 or 6, ratios are 0.986 and 0.992 respectively, therefore it is better to forecast the components and then aggregate the forecasts, as it leads to lower MSE while for h = 12 ratio becomes greater than one again, that is forecasting the aggregate directly is again more efficient. However, all ratios that are greater than 1 exceed unity by very little (the larger ratio is 1.023). Therefore, using the forecasts of the components cannot damage severely the forecasting accuracy.

In Table 19 we report the ratios of forecast MSEs of the forecast obtained by applying logs to the components and then aggregating the forecasts relative to the direct forecast MSEs of the aggregate.

Table 19: Ratios of forecast MSEs of disaggregate forecast in logs relative to linear forecast

Forecast horizon	$\left(\frac{MSE\left(y_{t+h\mid t}^{naidis}\right)}{MSE\left(y_{t+h\mid t}^{lin}\right)}\right)$
1	1.113
3	1.150
6	1.297
12	1.600

Now all ratios are again greater than 1, and some substantially so. Thus, applying logarithms to the disaggregate series and then aggregating the forecasts leads to higher MSEs. In other words, the forecast of the aggregate obtained that way is inferior to the direct forecast from the aggregated series. Furthermore, the situation gets worse as the forecast horizon increases.

If someone decides to apply logarithms, is it better to use them in the aggregate series or the components? To investigate this, ratios of forecast MSEs of the disaggregate forecasts based on logs relative to the aggregate forecast based on logs have to be calculated. The results are presented in Table 20. All entries in the table are greater than 1, that is $MSE\left(y_{t+h|t}^{naidis}\right) > MSE\left(y_{t+h|t}^{nai}\right)$ Therefore, it is beneficial in terms of MSE to apply logs directly on the aggregate series rather than to its components . The result holds for all forecast horizons.

Table 20: Ratios of forecast MSE of disaggregate forecast relative to naive forecast

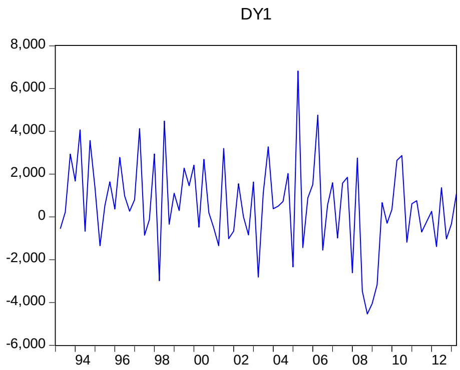
Forecast horizon	$\left(\frac{MSE\left(y_{t+h\|t}^{naidis}\right)}{MSE\left(y_{t+h\|t}^{nai}\right)}\right)$
1	1.044
3	1.038
6	1.025
12	1.017

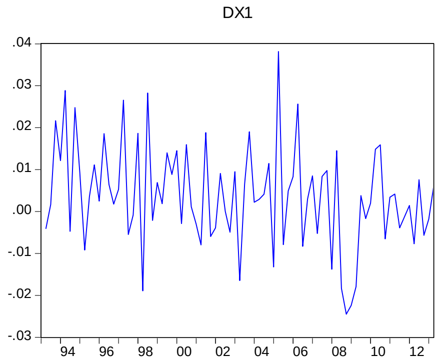
The overall conclusion of the aforementioned results is that using logarithms can be harmful for the forecasting accuracy, so it is preferable to forecast the aggregate directly. However, if anyone decides to apply logarithms, it is better to use them on the aggregated series directly.

4.2 Scandinavian GDP

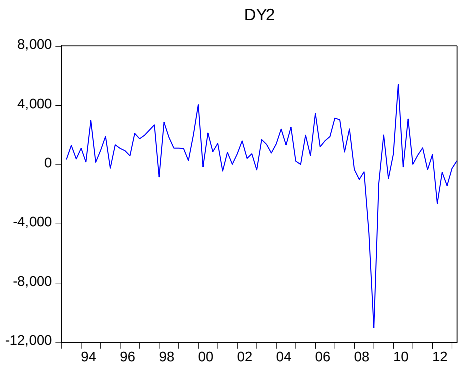
We now consider seasonally adjusted quarterly data of real GDP measured in millions of US dollars, in fixed PPPs for the Scandinavian countries (Denmark, Finland, Norway and Sweden), starting at 1993Q1 until 2013Q2. The data source is OECD. Scandinavian aggregated real GDP is computed as the sum of the four GDPs. First differences and first differences of logs of all the components and the aggregate are presented in Figure 2. Again, applying logs in the levels does not reduce variation in the volatility of the variables of interest.

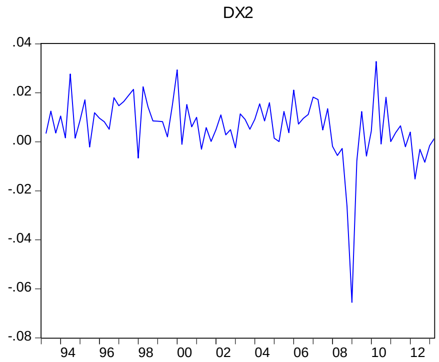
Denmark



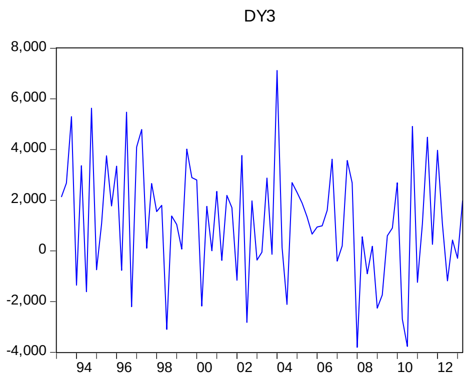


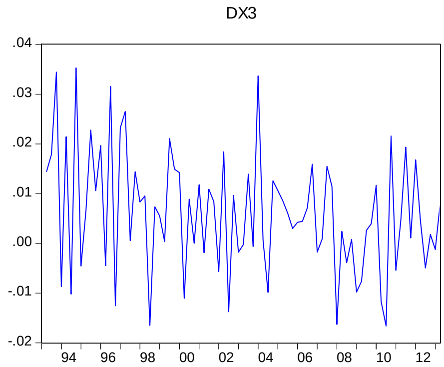
Finland



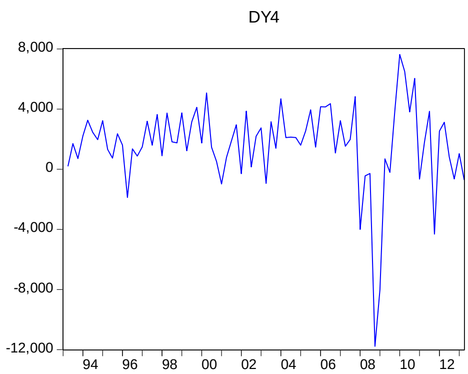


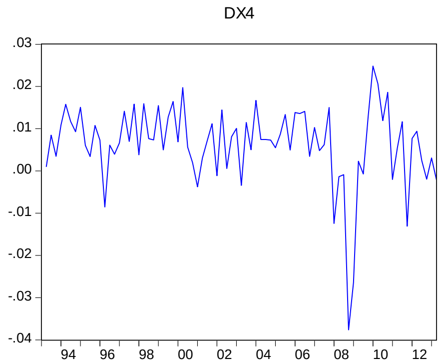
Norway





Sweden





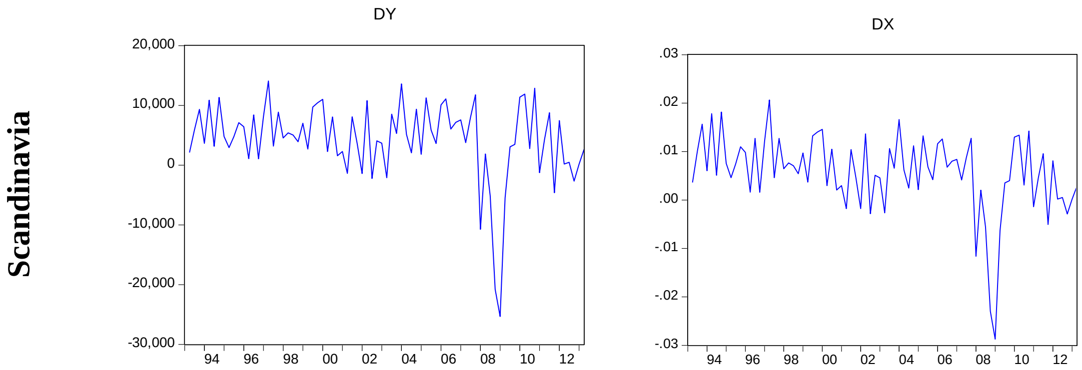


Figure 2: First differences and first differences of logs, where Y1, Y1, Y3, Y4 are the component series, Y is the aggregate and X1, X2, X3, X4, X are the respective logs

Augmented Dickey Fuller (ADF) tests are considered for all the components and the aggregate and the presence of a unit root cannot be rejected. ADF test for the first differences rejects non - stationarity, therefore all variables can be considered to be integrated of order 1, that is I(1). As a result, we fit AR(p) models to the first differences of the original variables and the variables in logs. The order p is determined using the SIC criterion with maximum lag order equal to 4. We compute the four different forecasts of the aggregated series that we presented in the simulation experiment for forecast horizons up to 12 periods.

The first comparison is between the forecast obtained by applying logs to the aggregated series and the forecast obtained by the aggregated series directly. For this purpose, ratios of the corresponding MSEs are computed and the results are reported in Table 21.

Table 21 : Ratios of forecast MSEs of the naive relative to the linear forecast

Forecast horizon	$\left(\frac{MSE\left(y_{t+h\mid t}^{nai}\right)}{MSE\left(y_{t+h\mid t}^{lin}\right)}\right)$
1	1.054
3	1.052
6	0.824
12	0.247

From the results it can be seen that for forecast horizons h = 1 or 3, ratios are greater than 1. In other words, $MSE\left(y_{t+h\mid t}^{nai}\right) > MSE\left(y_{t+h\mid t}^{lin}\right)$ and forecasting directly the aggregate leads to lower MSE, and as a consequence, more accurate forecasts. The situation changes for long forecast horizons, as for h = 6 or 12, where the ratio is less than 1, that is applying logs to the aggregate series and then obtaining the forecast for the aggregate is beneficial for forecasting accuracy. Especially for h = 12, the MSE of the forecast with logs is more than 4 times smaller than the MSE of the aggregate forecast.

The second comparison is between the forecast of the direct aggregate and the forecast obtained by aggregating the forecasts of the components. The ratios of the corresponding MSEs are reported in Table 22.

Table 22: Ratios of forecast MSEs of the disaggregate relative to the linear forecast

Forecast horizon	$\left(\dfrac{MSE\left(y_{t+h\mid t}^{dis}\right)}{MSE\left(y_{t+h\mid t}^{lin}\right)}\right)$
1	0.956
3	0.854
6	0.890
12	1.303

For forecast horizon h = 1, 3, 6 ratios are smaller than 1, that is aggregating the forecasts of the disaggregated series leads to lower MSE and thus is beneficial. For longer forecast horizons, ratios are larger than 1 and furthermore they increase as the forecast horizon increases. In other words, for long forecast horizons forecasting directly the aggregate may lead to lower MSE and thus is preferable.

The third comparison is between the forecast of the aggregate and the one obtained by aggregating the forecasts of the components after applying logs. In Table 23, we report ratios of forecast MSEs of the disaggregate relative to the linear forecast of the aggregate.

Table 23: Ratios of forecast MSEs of the disaggregate forecast based on logs relative to the linear forecast

Forecast horizon	$\left(\dfrac{MSE\left(y_{t+h\mid t}^{naidis}\right)}{MSE\left(y_{t+h\mid t}^{lin}\right)}\right)$
1	0.925
3	0.746
6	0.505
12	0.106

Now all entries are less than 1, that is $MSE\left(y_{t+h|t}^{naidis}\right) < MSE\left(y_{t+h|t}^{lin}\right)$. So forecasting the components based on logs and then aggregating the forecasts leads to lower MSEs and improves forecast accuracy for all forecast horizons. Furthermore, as the forecast horizon increases, the ratios decrease, therefore the difference between the MSEs gets larger and the superiority of the disaggregate forecast with logs increases. Especially for the 12 – step ahead forecast, the ratio is 0.106, that is the MSE of the disaggregate forecast is almost 10 times smaller than the MSE of the aggregate.

As it is obvious from the above, using logarithms improves forecast accuracy, at least for long forecast horizons. Now the question is whether it is better to apply logs on the aggregate series or on its components. For this reason, ratios of forecast MSEs of the disaggregate relative to the aggregate forecast based on logs are computed and reported in Table 24.

Table 24: Ratios of forecast MSEs of the disaggregate forecast based on logs relative to the naive forecast

Forecast horizon	$\left(\frac{MSE\left(y_{t+h\|t}^{naidis}\right)}{MSE\left(y_{t+h\|t}^{nai}\right)}\right)$
1	0.878
3	0.709
6	0.612
12	0.431

All entries in Table 24 are smaller than 1, and some substantially so. In other words, it is better to apply the logarithmic transformation on the disaggregate components and then aggregate the forecasts, as it leads to lower MSE and thus more accurate forecasts. The gains are even more substantial for longer forecast horizons where the ratios become smaller.

To conclude, applying logs to the variables of interest and then obtain the forecasts by exponentiating can be beneficial for all forecast horizons when the logs are applied to the components. If the logs are applied to the aggregated series, the benefits occur for forecast horizon h > 3. As for the aggregation, forecasting the components and then aggregating the forecasts can be beneficial for forecast horizon h < 9. For longer forecast horizons it is preferable to forecast the aggregate directly.

5. Conclusions

Applying logarithms in variables of interest for estimation and forecasting analysis is a common method in economics, as it appears to offer a more homogenous variance to the variables. This thesis studies whether forecasting the logarithm of a variable y_t and then converting it to a forecast of y_t leads to a more accurate forecast than forecasting directly the variable y_t. The study is extended to variables that are aggregates of a number of individual components. The forecast of the aggregate is compared to other three forecasts: one obtained by applying logs to the aggregated series and then transforming it through exponentiation to forecast of the aggregate, one obtained by aggregating the forecasts of the components, and one obtained by aggregating the forecasts of the components based on logs.

Two simulation experiments are carried. The results show that using logarithms is beneficial for forecasting, as long as it stabilizes the variance of the variable of interest. If it does not improve homogeneity, applying logs may be harmful for forecast precision. If the variable of interest is an aggregate, applying logs either on the aggregate or the disaggregated series is also beneficial for forecasting accuracy. However, the gains are more substantial if logs are applied on the aggregate. If logs are not applied at all, then aggregating the forecasts of the components is preferable than forecasting the aggregate directly only for large samples.

Two empirical applications are carried out using real economic aggregated variables, the real GDP of Australia and Scandinavia. The results from the Australian GDP show that the log transformation can be harmful for the forecasting accuracy, as it leads to higher MSEs while simply aggregating the forecasts of the components leads to ambiguous results. The results from Scandinavian GDP comparison tend to favor the log transformation. Using logs can be beneficial for the forecasting accuracy and the gains can be dramatic especially for long forecast horizons. When it comes to the aggregation effect, aggregating forecasts of the components is beneficial only for short forecast horizons. For long forecast horizons, it is better to forecast the aggregate directly. Therefore, forecast accuracy in practice depends on the dataset and should be treated with caution.

References

Arino M, Franses P. H (2000) Forecasting the levels of vector autoregressive log – transformed time series. *International Journal of Forecasting* vol. 16(1), p. 111 - 116

Bardsen G, Lütkepohl H (2011) Forecasting levels of log variables in vector autoregressions. *International Journal of Forecasting* vol. 27, p. 1108 – 1115

Benalal N, Diaz del Hoyo J L, Landau B, Roma M, Skudelny F (2004) To aggregate of not to aggregate? Euro area inflation forecasting. Working Paper Series No.374/July 2004, European Central Bank

Bermingham C, D'Agostino A (2011) Understanding and forecasting aggregate and disaggregate price dynamics. Working Paper Series No. 1365/August 2011, European Central Bank

Brüggemann R, Lütkepohl H (2013) Forecasting contemporaneous aggregates with stochastic aggregation weights. *International Journal of Forecasting* vol. 29, p. 60 – 68

Duarte C, Rua A (2005) Forecasting inflation through a bottom – up approach: the Portuguese case. Working Papers w200502, Economic Research Department, Banco de Portugal

Espasa A, Mayo I (2012) Forecasting aggregates and disaggregates with common features. Statistics and Econometric Series (05), Working Paper 11 – 08, Universidad Carlos III de Madrid

Espasa A, Senra E, Albacete R (2001) Forecasting inflation in the European Monetary Union: A disaggregated approach by countries and by sectors. Working Paper 01 – 37, Statistics and Econometrics Series 23, Universidad Carlos III de Madrid

Flavin T, Panopoulou E, Pantelidis T (2009) Forecasting growth and inflation in an enlarged Euro area. *Journal of Forecasting*, vol. 28(5), p. 405 – 425

Giacomini R, Granger C (2002) Aggregation of Space – Time Processes. Boston College Working

Papers in Economics 582, Boston College Department of Economics

Granger C, Newbold P (1976) Forecasting transformed series. *Journal of the Royal Statistical Society*, Series B, vol. 38(2), p. 189 – 203

Hendry D, Hubrich K (2006) Forecasting economic aggregates by disaggregates. Working Paper Series, No 589/ February 2006, European Central Bank

Hendry D. F, Hubrich K (2010) Combining disaggregate forecasts or combining disaggregate Information to forecast an aggregate. Working Paper Series, No 1155/ February 2010, European Central Bank

Hubrich K (2004) Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy? Computing in Economics and Finance 2004 230, Society for Computational Economics

Kohn R (1982) When is an aggregate of a time series efficiently forecast by its past? *Journal of Econometrics* vol. 18, p. 337 – 349

Lütkepohl H (1984) Linear transformations of vector ARMA processes. *Journal of Econometrics* vol. 26, p. 283 – 293

Lütkepohl H (1986) Forecasting temporally aggregated vector ARMA processes. *Journal of Forecasting*, vol. 5, p. 85 – 95

Lütkepohl H (1987) *Forecasting aggregated vector ARMA processes*. Springer, Berlin

Lütkepohl H (2010) Forecasting nonlinear aggregates and aggregates with time – varying weights. EUI Working Papers ECO 2010/11, European University Institute

Lütkepohl H, Xu F (2011) Forecasting annual inflation with seasonal monthly data: Using levels versus logs of the underlying price index. Journal of Time Series Econometrics, De Gruyter, vol.3(1), p. 1 – 23, February

Lütkepohl H, Xu F (2012) The role of the log transformation in forecasting economic variables. *Empirical Economics* vol. 42, p. 619 – 638

Marcellino M, Stock J , Watson M (2001) Macroeconomic forecasting in the Euro area: Country specific versus area – wide information. Working Paper No. 201, IGIER, Bocconi University

Mayr J, Ulbricht D (2007) Log versus level in VAR forecasting: 16 million empirical answers – expect the unexpected. Ifo Working Paper Series, Ifo Working Paper No. 42, Ifo Institute for Economic Research at the University of Munich

Moser G, Rumler F, Scharler J (2004) Forecasting Austrian inflation. Working Paper 91, Oesterreichische Nationalbank

Sbrana G (2012) Forecasting aggregated moving average processes with an application to the Euro area real interest rate. *Journal of Forecasting* vol. 31, p. 85 – 98

Sbrana G, Sivestrini A (2009) What do we know about comparing aggregate and disaggregate forecasts? CORE Discussion Paper 2009/20, Universite catholique de Louvain

Tiao G.C, Guttman I (1980) Forecasting contemporal aggregates of multiple time series. *Journal of Econometrics* vol. 12, p. 219 – 230

Wei W, Abraham B (1981) Forecasting contemporal time series aggregates. *Communications in Statistics – Theory and Methods* vol. A10(13), p. 1335 - 1344

Appendix

Table 1a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 40 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.957	0.952	0.945	0.940	0.935	0.937	0.907	0.910	0.918	0.873	0.895	0.864	0.858	0.842	0.853	0.895	0.940	0.980	1.022
3	0.892	0.891	0.878	0.870	0.844	0.845	0.782	0.801	0.778	0.749	0.735	0.695	0.649	0.672	0.642	0.692	0.649	0.831	1.217
6	0.836	0.811	0.773	0.764	0.715	0.725	0.677	0.655	0.644	0.600	0.580	0.528	0.517	0.552	0.412	0.460	0.457	0.641	0.960
12	0.703	0.673	0.615	0.595	0.566	0.558	0.510	0.492	0.481	0.443	0.443	0.385	0.362	0.386	0.111	0.184	0.097	0.384	0.661

Table 1b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 40 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.640	0.621	0.606	0.583	0.607	0.584	0.599	0.593	0.613	0.612	0.629	0.643	0.678	0.698	0.741	0.789	0.785	0.865	0.933
3	0.448	0.405	0.369	0.330	0.316	0.293	0.278	0.271	0.270	0.290	0.309	0.336	0.367	0.417	0.471	0.523	0.576	0.671	0.846
6	0.300	0.244	0.206	0.189	0.178	0.162	0.151	0.143	0.142	0.150	0.157	0.174	0.198	0.237	0.276	0.328	0.388	0.479	0.782
12	0.154	0.129	0.110	0.103	0.097	0.089	0.078	0.077	0.075	0.078	0.078	0.083	0.095	0.117	0.135	0.158	0.206	0.296	0.688

Table 2a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 80 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.853	0.856	0.844	0.833	0.820	0.807	0.806	0.804	0.793	0.797	0.786	0.781	0.841	0.786	0.781	0.804	0.823	0.684	1.263
3	0.756	0.732	0.688	0.665	0.638	0.626	0.597	0.581	0.569	0.561	0.560	0.560	0.606	0.563	0.574	0.606	0.623	0.470	0.845
6	0.612	0.566	0.520	0.484	0.477	0.465	0.430	0.404	0.403	0.410	0.399	0.402	0.439	0.397	0.388	0.408	0.409	0.467	0.517
12	0.429	0.374	0.363	0.320	0.315	0.316	0.289	0.267	0.274	0.264	0.273	0.269	0.293	0.223	0.161	0.175	0.128	0.300	0.306

Table 2b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 80 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.384	0.404	0.410	0.426	0.441	0.451	0.465	0.502	0.535	0.547	0.582	0.619	0.671	0.677	0.722	0.779	0.826	0.828	0.762
3	0.228	0.210	0.184	0.185	0.172	0.175	0.181	0.206	0.227	0.266	0.290	0.338	0.383	0.407	0.457	0.527	0.610	0.671	0.716
6	0.124	0.109	0.097	0.092	0.085	0.086	0.085	0.094	0.102	0.117	0.143	0.179	0.207	0.243	0.272	0.347	0.420	0.501	0.930
12	0.064	0.056	0.050	0.049	0.045	0.044	0.044	0.044	0.047	0.050	0.064	0.080	0.095	0.119	0.141	0.187	0.218	0.295	1.093

Table 3a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 160 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.739	0.728	0.709	0.716	0.730	0.726	0.734	0.709	0.722	0.765	0.748	0.744	0.731	0.727	0.883	1.383	0.722	0.702	0.946
3	0.583	0.537	0.520	0.510	0.484	0.455	0.470	0.442	0.440	0.496	0.492	0.502	0.506	0.486	0.757	1.449	0.515	0.532	0.655
6	0.417	0.381	0.345	0.339	0.309	0.310	0.300	0.294	0.274	0.318	0.331	0.346	0.380	0.363	0.471	0.759	0.300	0.392	1.389
12	0.255	0.233	0.208	0.210	0.198	0.210	0.192	0.186	0.169	0.195	0.209	0.212	0.248	0.219	0.249	0.158	0.176	0.127	1.212

Table 3b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 160 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.262	0.277	0.309	0.330	0.374	0.385	0.430	0.473	0.500	0.526	0.580	0.617	0.649	0.697	0.719	0.760	0.807	0.870	0.914
3	0.134	0.125	0.121	0.120	0.130	0.143	0.162	0.198	0.221	0.254	0.290	0.328	0.367	0.420	0.454	0.528	0.573	0.717	0.892
6	0.067	0.059	0.054	0.054	0.055	0.059	0.068	0.081	0.099	0.120	0.151	0.175	0.211	0.241	0.279	0.343	0.394	0.538	0.779
12	0.033	0.030	0.028	0.027	0.027	0.027	0.030	0.034	0.040	0.049	0.062	0.078	0.101	0.120	0.142	0.176	0.214	0.327	0.664

Table 4a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 320 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameters																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.645	0.637	0.633	0.645	0.635	0.700	0.654	0.658	0.682	0.723	0.728	0.704	0.726	0.745	0.583	0.957	0.837	1.111	0.587
3	0.462	0.443	0.400	0.421	0.380	0.404	0.381	0.390	0.417	0.433	0.458	0.463	0.540	0.546	0.467	0.853	0.730	1.041	0.571
6	0.303	0.271	0.244	0.248	0.238	0.241	0.243	0.229	0.251	0.286	0.287	0.312	0.341	0.360	0.266	0.555	0.461	0.621	0.578
12	0.181	0.165	0.147	0.153	0.152	0.147	0.150	0.142	0.159	0.169	0.179	0.191	0.204	0.241	0.081	0.272	0.276	0.297	0.486

Table 4b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 320 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.196	0.226	0.258	0.297	0.340	0.375	0.402	0.461	0.486	0.523	0.576	0.596	0.655	0.684	0.697	0.754	0.810	0.945	0.686
3	0.093	0.091	0.097	0.100	0.120	0.132	0.161	0.188	0.219	0.252	0.304	0.321	0.366	0.420	0.440	0.512	0.560	0.674	0.453
6	0.043	0.040	0.040	0.042	0.049	0.054	0.066	0.079	0.097	0.119	0.158	0.173	0.207	0.249	0.272	0.329	0.371	0.442	0.417
12	0.022	0.020	0.020	0.020	0.021	0.023	0.026	0.031	0.038	0.047	0.065	0.080	0.101	0.122	0.134	0.177	0.204	0.289	0.334

Table 5a: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 40 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.001	1.001	1.001	1.001	1.001	1.002	1.001	1.001	1.002	1.001	1.002	1.002	1.002	1.003	1.004	1.003	1.009	1.006	1.009
3	1.001	1.001	1.002	1.003	1.002	1.004	1.003	1.004	1.006	1.005	1.005	1.006	1.007	1.012	1.014	1.016	1.029	1.027	1.058
6	1.003	1.003	1.004	1.005	1.004	1.008	1.007	1.007	1.012	1.009	1.013	1.012	1.020	1.027	1.034	1.042	1.072	1.080	1.126
12	1.006	1.008	1.009	1.009	1.009	1.013	1.013	1.015	1.019	1.018	1.027	1.027	1.040	1.056	1.075	1.094	1.182	1.351	1.364

Table 5b: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 40 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.002
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001	1.001	1.001	1.001	1.001	1.002	1.003	1.004	1.008
6	1.001	1.000	1.000	1.000	1.000	1.001	1.000	1.001	1.001	1.001	1.002	1.001	1.002	1.002	1.003	1.004	1.006	1.011	1.019
12	1.001	1.001	1.000	1.001	1.001	1.001	1.001	1.002	1.001	1.003	1.003	1.003	1.003	1.004	1.006	1.009	1.015	1.027	1.051

Table 6a: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 80 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.002	1.001	1.002	1.001	1.001	1.004	1.003	1.004	1.005	1.007	1.006	1.028
3	1.002	1.002	1.002	1.002	1.003	1.003	1.003	1.004	1.004	1.006	1.006	1.006	1.011	1.008	1.017	1.017	1.035	1.040	1.070
6	1.003	1.004	1.003	1.003	1.005	1.006	1.006	1.007	1.010	1.013	1.013	1.011	1.020	1.020	1.039	1.039	1.078	1.134	1.161
12	1.007	1.006	1.008	1.006	1.008	1.012	1.012	1.015	1.018	1.021	1.029	1.025	1.038	1.050	1.075	1.090	1.181	1.393	1.452

Table 6b: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 80 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.000	1.000	1.001	1.001
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.000	1.001	1.001	1.001	1.002	1.002	1.002	1.004	1.016
6	1.000	1.000	1.001	1.001	1.001	1.001	1.000	1.001	1.001	1.001	1.002	1.002	1.002	1.002	1.004	1.005	1.006	1.011	1.045
12	1.000	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.002	1.002	1.003	1.003	1.004	1.005	1.008	1.010	1.015	1.029	1.088

Table 7a: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 160 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameters																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.001	1.001	1.000	1.001	1.002	1.001	1.002	1.001	1.001	1.003	1.001	1.003	1.002	1.004	1.007	1.012	1.006	1.005	1.034
3	1.002	1.002	1.002	1.003	1.003	1.002	1.005	1.003	1.003	1.007	1.005	1.008	1.011	1.014	1.021	1.032	1.020	1.056	1.078
6	1.002	1.004	1.004	1.006	1.006	1.005	1.008	1.008	1.006	1.010	1.013	1.017	1.020	1.031	1.035	1.077	1.051	1.124	1.200
12	1.006	1.007	1.007	1.009	1.010	1.013	1.013	1.015	1.014	1.018	1.024	1.030	1.041	1.057	1.063	1.185	1.083	1.291	1.600

Table 7b: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 160 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameters																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001	1.001	1.002
3	1.000	1.000	1.001	1.000	1.000	1.000	1.000	1.001	1.000	1.000	1.001	1.001	1.000	1.002	1.002	1.002	1.003	1.006	1.001
6	1.000	1.001	1.000	1.000	1.000	1.000	1.001	1.001	1.000	1.001	1.001	1.001	1.001	1.003	1.003	1.005	1.007	1.014	1.012
12	1.000	1.001	1.001	1.000	1.001	1.001	1.002	1.002	1.001	1.001	1.002	1.003	1.003	1.007	1.009	1.012	1.017	1.032	1.061

Table 8a: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 320 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.001	1.001	1.001	1.002	1.000	1.004	1.000	1.002	1.001	1.003	1.002	1.001	1.003	0.998	1.013	1.011	1.013	1.016	0.982
3	1.002	1.002	1.002	1.003	1.002	1.004	1.003	1.004	1.003	1.005	1.006	1.007	1.015	1.002	1.029	1.031	1.054	1.060	0.931
6	1.004	1.003	1.002	1.005	1.005	1.006	1.006	1.007	1.008	1.012	1.015	1.017	1.023	1.014	1.052	1.079	1.091	1.132	0.924
12	1.007	1.007	1.005	1.009	1.011	1.013	1.011	1.014	1.017	1.023	1.028	1.039	1.040	1.053	1.105	1.139	1.188	1.213	1.212

Table 8b: Forecast MSEs of optimal relative to naive forecast for DGP $\Delta x_t = c + a\Delta x_{t-1} + e_t$, T = 320 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter																		
h	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.000	1.000	1.001	1.002	0.997
3	1.000	1.000	1.000	1.000	1.001	1.000	1.000	1.000	1.000	1.000	1.001	1.000	1.000	1.002	1.001	1.003	1.003	1.006	1.004
6	1.000	1.000	1.000	1.000	1.001	1.000	1.001	1.000	1.001	1.000	1.001	1.001	1.002	1.004	1.003	1.005	1.010	1.018	1.030
12	1.000	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.003	1.003	1.005	1.007	1.007	1.011	1.022	1.042	1.112

Table 9a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 40 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	1.135	1.126	1.127	1.127	1.119	1.131	1.143	1.124	1.132	1.132	1.138	1.129	1.140	1.161	1.145
3	1.296	1.365	1.392	1.424	1.405	1.419	1.436	1.385	1.403	1.401	1.406	1.395	1.404	1.395	1.375
6	1.675	1.829	1.908	1.969	1.922	1.939	2.021	1.881	1.893	1.890	1.892	1.897	1.863	1.763	1.727
12	2.762	3.239	3.205	3.240	3.273	3.316	3.314	3.138	3.196	3.168	3.149	3.132	2.997	2.801	2.517

Table 9b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 40 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	1.882	1.832	1.782	1.774	1.769	1.752	1.679	1.666	1.611	1.605	1.513	1.485	1.427	1.314	1.124
3	3.093	3.346	3.618	3.677	3.761	3.693	3.622	3.605	3.450	3.106	2.943	2.717	2.407	2.005	1.360
6	5.854	6.743	7.118	7.489	7.350	7.437	7.043	6.945	6.738	6.276	5.946	5.396	4.551	3.404	1.854
12	14.153	16.345	16.385	17.084	17.543	17.171	16.087	15.854	15.176	14.874	13.990	12.942	11.392	7.701	3.262

Table 10a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 80 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	1.315	1.299	1.294	1.285	1.296	1.310	1.286	1.265	1.279	1.249	1.293	1.285	1.269	1.249	1.164
3	1.734	1.828	1.891	1.939	1.964	1.988	1.975	1.934	1.905	1.875	1.875	1.851	1.768	1.678	1.418
6	2.500	2.873	2.982	2.988	3.058	3.047	3.177	3.058	3.003	2.861	2.823	2.837	2.684	2.446	1.802
12	4.884	5.521	5.705	5.651	5.834	5.815	5.747	5.635	5.810	5.403	5.214	5.373	5.107	4.480	2.744

Table 10b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 80 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	3.421	3.199	2.899	2.757	2.637	2.485	2.349	2.204	2.056	1.909	1.736	1.650	1.529	1.238	1.093
3	6.796	7.940	8.273	7.809	7.843	7.400	6.926	6.262	5.781	4.980	4.234	3.553	2.976	1.808	1.273
6	14.336	17.268	17.859	17.211	16.882	16.232	15.600	14.746	13.966	12.200	10.677	8.334	6.580	3.116	1.617
12	33.765	40.229	40.447	39.426	38.744	38.790	37.228	36.573	34.148	31.273	29.582	23.924	18.938	8.040	2.545

Table 11a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 160 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameter														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	1.716	1.659	1.680	1.615	1.672	1.618	1.618	1.605	1.544	1.541	1.496	1.502	1.434	1.319	1.129
3	2.603	2.815	3.101	3.027	3.196	3.095	3.061	3.078	2.961	2.880	2.734	2.637	2.392	1.956	1.338
6	4.212	5.122	5.487	5.470	5.742	5.645	5.686	5.358	5.404	5.204	4.888	4.813	4.206	3.117	1.714
12	9.177	10.708	11.001	10.739	11.332	11.133	11.627	10.703	10.716	10.460	10.385	9.953	9.177	6.496	2.718

Table 11b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 160 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameter														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	6.762	5.617	4.533	4.086	3.695	3.240	2.853	2.630	2.281	2.083	1.818	1.659	1.502	1.267	1.114
3	14.517	16.691	17.181	16.135	15.121	13.316	11.100	9.644	8.097	6.303	4.884	3.852	3.019	1.871	1.354
6	31.964	39.035	40.125	38.463	37.011	34.024	30.093	26.730	22.270	18.338	13.646	10.219	7.044	3.196	1.763
12	76.535	86.582	91.246	88.706	86.443	82.657	75.534	70.478	64.418	56.852	44.158	33.720	23.737	8.105	2.890

Table 12a: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 320 and $\sigma_e^2 = 0.001$

Forecast horizon	AR parameters														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	2.583	2.460	2.355	2.438	2.220	2.236	2.092	2.081	2.015	1.916	1.844	1.685	1.615	1.387	1.148
3	4.487	5.064	5.442	5.454	5.464	5.447	5.124	4.901	4.794	4.468	4.197	3.629	3.216	2.276	1.370
6	8.413	9.449	10.699	10.492	10.420	10.723	10.128	9.355	9.644	9.343	8.832	7.742	6.743	4.268	1.776
12	17.625	20.193	21.362	21.502	21.164	21.342	20.470	20.203	19.647	20.085	19.614	17.815	15.959	9.716	2.833

Table 12b: Forecast MSEs of naive relative to linear forecast for DGP $\Delta y_t = c + a\Delta y_{t-1} + e_t$, T = 320 and $\sigma_e^2 = 0.0001$

Forecast horizon	AR parameters														
h	-0.9	-0.7	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9
1	12.500	8.915	6.279	5.271	4.395	3.710	3.338	2.809	2.519	2.140	1.884	1.680	1.493	1.261	1.159
3	29.933	31.354	31.147	28.500	23.567	19.923	16.064	12.426	9.454	7.138	5.420	3.964	2.925	1.888	1.469
6	69.884	80.672	74.654	71.565	65.599	57.682	49.553	40.881	30.365	23.788	16.893	11.316	7.131	3.293	2.043
12	157.303	175.688	178.487	170.782	164.891	149.059	137.339	124.387	101.236	82.709	61.804	42.242	26.060	8.540	3.766

Table 13: Ratios of forecast MSEs of the naive relative to the linear forecast of the aggregate $\left(\frac{MSE\left(y_{t+h|t}^{nai}\right)}{MSE\left(y_{t+h|t}^{lin}\right)}\right)$

Sample size	Forecast horizon											
T	1	2	3	4	5	6	7	8	9	10	11	12
40	0.990	0.975	0.958	0.935	0.923	0.913	0.904	0.887	0.876	0.866	0.856	0.840
80	0.912	0.855	0.809	0.759	0.726	0.696	0.663	0.643	0.622	0.589	0.569	0.551
160	0.864	0.774	0.701	0.638	0.589	0.545	0.492	0.455	0.417	0.400	0.376	0.354
320	0.844	0.948	0.884	0.808	0.722	0.687	0.639	0.550	0.517	0.503	0.465	0.443

Table 14: Ratios of forecast MSEs of the disaggregate relative to the linear forecast of the aggregate $\left(\frac{MSE\left(y_{t+h|t}^{dis}\right)}{MSE\left(y_{t+h|t}^{lin}\right)}\right)$

Sample size	Forecast horizon											
T	1	2	3	4	5	6	7	8	9	10	11	12
40	1.015	1.007	1.002	1.006	1.009	1.011	1.007	1.006	1.005	1.005	1.004	1.006
80	1.018	1.024	1.035	1.047	1.050	1.063	1.079	1.082	1.092	1.095	1.107	1.115
160	0.997	1.000	0.999	0.997	0.995	0.993	0.992	0.994	0.997	0.998	0.998	0.999
320	1.002	0.998	0.971	0.955	0.956	0.963	0.959	0.950	0.948	0.951	0.949	0.950

Table 15: Ratios of forecast MSEs of the disaggregate forecast based on logs relative to the the linear forecast of the aggregate $\left(\frac{MSE\left(y_{t+h|t}^{naidis}\right)}{MSE\left(y_{t+h|t}^{lin}\right)}\right)$

Sample size	Forecast horizon											
T	1	2	3	4	5	6	7	8	9	10	11	12
40	1.009	0.987	0.971	0.951	0.942	0.939	0.936	0.918	0.910	0.903	0.899	0.884
80	0.931	0.874	0.833	0.787	0.756	0.728	0.697	0.679	0.665	0.627	0.612	0.599
160	0.871	0.779	0.704	0.638	0.586	0.545	0.492	0.457	0.419	0.408	0.386	0.361
320	0.820	0.993	0.953	0.855	0.794	0.771	0.733	0.638	0.600	0.554	0.523	0.504

Table 16: Ratios of forecast MSEs of the disaggregate forecast based on logs relative to the naive forecast of the aggregate $\left(\frac{MSE\left(y_{t+h|t}^{naidis}\right)}{MSE\left(y_{t+h|t}^{nai}\right)}\right)$

Sample size	Forecast horizon											
T	1	2	3	4	5	6	7	8	9	10	11	12
40	1.019	1.012	1.014	1.017	1.021	1.029	1.035	1.035	1.039	1.043	1.051	1.053
80	1.020	1.023	1.030	1.038	1.041	1.046	1.051	1.057	1.069	1.064	1.075	1.087
160	1.008	1.008	1.004	0.999	0.994	1.000	1.001	1.005	1.004	1.022	1.024	1.020
320	0.972	1.048	1.078	1.058	1.100	1.124	1.148	1.160	1.160	1.101	1.124	1.140

Table 17: MSE ratios (Australia)

Forecast horizon	$\left(\frac{MSE\left(y_{t+h\|t}^{nai}\right)}{MSE\left(y_{t+h\|t}^{lin}\right)}\right)$	$\left(\frac{MSE\left(y_{t+h\|t}^{dis}\right)}{MSE\left(y_{t+h\|t}^{lin}\right)}\right)$	$\left(\frac{MSE\left(y_{t+h\|t}^{naidis}\right)}{MSE\left(y_{t+h\|t}^{lin}\right)}\right)$	$\left(\frac{MSE\left(y_{t+h\|t}^{naidis}\right)}{MSE\left(y_{t+h\|t}^{nai}\right)}\right)$
1	1.067	1.023	1.113	1.044
2	1.093	1.011	1.146	1.048
3	1.108	0.986	1.150	1.038
4	1.158	0.970	1.186	1.024
5	1.222	0.977	1.252	1.024
6	1.265	0.992	1.297	1.025
7	1.312	0.997	1.336	1.018
8	1.352	1.001	1.377	1.019
9	1.418	0.999	1.447	1.020
10	1.486	1.001	1.520	1.023
11	1.522	1.004	1.553	1.021
12	1.573	1.007	1.600	1.017

Table 18: MSE ratios (Scandinavia)

Forecast horizon	$\left(\frac{MSE\left(y_{t+h\|t}^{nai}\right)}{MSE\left(y_{t+h\|t}^{lin}\right)}\right)$	$\left(\frac{MSE\left(y_{t+h\|t}^{dis}\right)}{MSE\left(y_{t+h\|t}^{lin}\right)}\right)$	$\left(\frac{MSE\left(y_{t+h\|t}^{naidis}\right)}{MSE\left(y_{t+h\|t}^{lin}\right)}\right)$	$\left(\frac{MSE\left(y_{t+h\|t}^{naidis}\right)}{MSE\left(y_{t+h\|t}^{nai}\right)}\right)$
1	1.054	0.956	0.925	0.878
2	1.082	0.880	0.823	0.761
3	1.052	0.854	0.746	0.709
4	0.994	0.850	0.663	0.667
5	0.906	0.866	0.587	0.648
6	0.824	0.890	0.505	0.612
7	0.711	0.930	0.420	0.590
8	0.607	0.981	0.338	0.556
9	0.504	1.043	0.262	0.519
10	0.407	1.117	0.200	0.492
11	0.321	1.204	0.147	0.457
12	0.247	1.303	0.106	0.431