

Forecasting Industrial Production Using Aggregated and Disaggregated Series

Subjects: **Economics**

Contributor: Diogo de Prince , Emerson Fernandes Marçal , Pedro L. Valls Pereira

Researchers whether using a disaggregated series or combining an aggregated and disaggregated series improves the forecasting of the aggregated series compared to using the aggregated series alone. Researchers used econometric techniques, such as the weighted lag adaptive least absolute shrinkage and selection operator, and Exponential Triple Smoothing (ETS), as well as the Autometrics algorithm to forecast industrial production in Brazil one to twelve months ahead.

industrial production

forecasting

model selection

1. Introduction

Economic agents make decisions based on their views on the present state of the economy and their expectations for the future. The general levels of output, employment, interest rates, exchange rates, and inflation are key economic indicators that help to diagnose a country's economic situation. Therefore, the proposition and evaluation concerning the ability of econometric models to forecast a country's economic reality introduce benefits that create better guides for economic agents and policymakers.

One of the main macroeconomic indicators of an economy is the gross domestic product (GDP), which is a proxy for a country's economic performance. Researchers use industrial production as a proxy for the GDP since the monthly industrial production index is of higher frequency than the GDP. Moreover, the industrial production index is released with a lag of one month, which is smaller than that of the GDP, which has a delayed release of more than two months.

Researchers address whether using a disaggregated series or combining an aggregated and disaggregated series improves the forecasting accuracy of the aggregated series compared to using the aggregated series alone for the industrial production in Brazil series. Disaggregated data refer to the decomposition of the main variable into several sub-components, which have different weights for the aggregated series. Researchers obtained a forecast of these sub-components individually and then researchers grouped the forecasts of these sub-components to estimate the forecast of the aggregated series. This alternative could increase the accuracy of the forecast; researchers modeled the sub-components by taking their characteristics into account. Researchers used this alternative in the present work to understand if there was a reduction in the forecast error of the aggregate series by estimating a model for each sub-component.

The literature addresses the accuracy of using disaggregated or aggregated data for forecasting. According to [Lütkepohl \(1987\)](#), the forecast using disaggregation is theoretically optimal if the disaggregated series is uncorrelated; the author suggests using disaggregation if the correlation between the disaggregated series is not strong. Some examples of contributions to the theoretical literature on aggregate or disaggregate forecasting include [Lütkepohl \(1984, 1987\)](#); [Granger \(1987\)](#); [Pesaran et al. \(1989\)](#); [Van Garderen et al. \(2000\)](#); and [Giacomini and Granger \(2004\)](#). The following questions arise: Does aggregating a disaggregated forecast improve the accuracy of the aggregate forecast? One alternative is to estimate using only the lagged aggregate variable to forecast the aggregate series. [Giacomini \(2015\)](#) points out that the results of the empirical literature are mixed, but that disaggregation can improve the forecast accuracy of the aggregate variable. Another alternative is to combine the disaggregate and aggregate series and select the relevant variables to forecast the aggregate series. [Hendry and Hubrich \(2011\)](#) suggest this as a promising direction when using model selection procedures, even though the authors developed a dynamic factor model to consider the disaggregation and did not develop a selection procedure.

Researchers' goal was to determine whether forecasting the disaggregated components of industrial production in Brazil or combining these components with the aggregate series, improve the forecast accuracy of Brazil's aggregate industrial production compared to using only the lagged aggregate variable. Researchers analyzed Brazil as the 9th GDP in dollars based on 2019 World Bank data. In addition, Brazil is an emerging economy, so it has a more volatile business cycle than advanced countries, a stylized fact in the literature as seen in [Aguilar and Gopinath \(2007\)](#), and in [Kohn et al. \(2021\)](#), among others. This higher volatility can lead to difficulty in forecasting Brazilian economic activity, another motivator for researchers' research.

Researchers do not know of any other articles that address the contributions of disaggregated data from the weighted lag adaptive least absolute shrinkage and selection operator (WLadaLASSO) methodology or from exponential triple smoothing (ETS), selecting the most appropriate model or the relevant variables from the combination of a disaggregate and aggregate series to forecast industrial production. Only [Bulligan et al. \(2010\)](#) analyzed the contributions of disaggregated data to forecast industrial production, and researchers intend to fill this gap. The topic of disaggregation or aggregation in forecasting is most commonly studied for the inflation and GDP series, such as [Espasa et al. \(2002\)](#); [Marcellino et al. \(2003\)](#); [Hubrich \(2005\)](#); [Carlo and Marçal \(2016\)](#); and [Heinisch and Scheufele \(2018\)](#). Additionally, researchers analyzed the forecast accuracies of the models based on the multi-horizon superior predictive ability method developed by [Quaedvlieg \(2021\)](#) by combining different horizons, which is different from other forecast comparison procedures that focus on the model performances of the models for each horizon separately. [Quaedvlieg \(2021\)](#) developed the average multi-horizon superior predictive ability (aSPA) and uniform multi-horizon superior predictive ability (uSPA) tests to compare a multi-horizon forecast. Using monthly data from January 2002 to February 2020, researchers selected the best model for a rolling window of 100 fixed observations and evaluated the forecast for industrial production in Brazil one to twelve months ahead. Researchers used 91 rolling windows. Researchers considered the first-order autoregressive model (AR(1)), AR(1) with time-varying parameters (TVP-AR(1)), the thirteenth-order autoregressive model (AR(13)), and the unobserved components with stochastic volatility (UC-SV) estimated based on [Barnett et al. \(2014\)](#) as naive models. Researchers also analyzed the following methods for selecting the

best model: ETS based on [Hyndman et al. \(2002, 2008\)](#) and [Hyndman and Khandakar \(2008\)](#), the least absolute shrinkage and selection operator (LASSO), adaptive LASSO (adaLASSO), the WLadaLASSO, and the Autometrics algorithm. Researchers used the LASSO and its variants to select the lags from the fifteenth-order autoregressive model (AR(15)). Additionally, researchers considered the Autometrics algorithm that selects the lags from an AR(15) and the dummy variables for outliers or breaks in the sample. In addition, researchers combined the disaggregated and aggregated series in the model to forecast the general industrial production. To reduce the dimensionality of this model with the combination, researchers adopted the LASSO and adaLASSO procedures, and the Autometrics algorithm. Researchers compared the forecasting performance between the models based on the mean square error (MSE), the modified [Diebold and Mariano \(1995\)](#) test (henceforth, the MDM test), the model confidence set (MCS) procedure from [Hansen et al. \(2011\)](#), the forecast encompassing test from [Harvey et al. \(1998\)](#), and the multi-horizon superior predictive ability from [Quaedvlieg \(2021\)](#).

Researchers' MSE results point to the ETS model having a better forecasting accuracy for industrial production in Brazil compared to other models. The disaggregated ETS model is the ETS model for each disaggregated series. The disaggregated ETS model leads to the lowest MSE among all of the models for all the forecast horizons, except for those that are one and two months ahead. For the forecasts that are one and two months ahead, the aggregate ETS model has a lower MSE, and there is little difference compared to the disaggregated ETS model. The aggregated ETS model is the ETS model using the lagged aggregated series as covariates. The disaggregated ETS model also has a lower MSE than the forecast of the combination of the aggregated and disaggregated series. This result is similar to that of [Faust and Wright \(2013\)](#), who determined that the combination of the disaggregated and aggregated series does not lead to a better forecast compared to aggregating the disaggregated forecasts; however, their study focused on the United States (US) consumer price index (CPI). Researchers' results are in the opposite direction of the results by [Hendry and Hubrich \(2011\)](#) and [Weber and Zika \(2016\)](#). To analyze whether there was better statistical performance, researchers used the ETS with disaggregated data as a benchmark in the MDM test. The disaggregated ETS model presents a better forecast performance compared to the naive models (AR(1), AR(13), TVP-AR(1), UC-SV), LASSO, and variants, and the Autometrics algorithm, considering aggregated and disaggregated data (or a combination of both). Only the aggregated ETS model has equal predictive accuracy to the disaggregated ETS model for the forecast horizons of one to five, seven, ten, and twelve months ahead based on the MDM test. The set of "best" models for the most forecast horizons includes only the disaggregated and aggregated ETS models with 90% probability according to the MCS. In 2 of the 12 forecasting horizons, the MCS only has the disaggregated ETS model. Researchers also used the forecast encompassing test. Results showed that the optimal combination forecast only incorporated forecasts from the disaggregated ETS model and the aggregated ETS model. The disaggregated ETS forecast was the only model to be considered in the optimal combination forecast of industrial production for 10 horizons among the 12 analyzed, comparable to the aggregated ETS model. Aggregated ETS does not contain information that is useful for forecasting industrial production in Brazil beyond the information already found in the disaggregated ETS between two and twelve months ahead. When researchers analyzed the 12 horizons together, researchers rejected the null hypothesis of equal predictability for all of the models compared to the disaggregated ETS by the uSPA and aSPA tests at 5% statistical significance. In short, researchers determined that the ETS

model presents the best forecast performance comparatively, which is a result similar to that of [Elliott and Timmermann \(2008\)](#). The disaggregated ETS is superior after 6 horizons when compared to the aggregated ETS based on the aSPA test. The aggregated ETS only introduces relevant information to forecast industrial production for one period ahead compared to the disaggregated ETS according to the forecast encompassing test, which indicates the superiority of disaggregated information for industrial production, which is in line with [Bulligan et al. \(2010\)](#).

2. Aggregating the disaggregated forecasts, only modeling the aggregate variable, and combining the aggregated and disaggregated series

This section discusses the differences in the forecast accuracy in three scenarios—aggregating the disaggregated forecasts, only modeling the aggregate variable, and combining the aggregated and disaggregated series. [Bulligan et al. \(2010\)](#) analyzed the forecasting performance of industrial production models in Italy with forecast horizons that ranged from 1 to 3 months ahead. They determined that disaggregated models have better forecast performance based on the root of MSE. There are not many analyses in the literature that differentiate between the use of the disaggregated and aggregated series to forecast the aggregated series of industrial production. As such, researchers have to fill in this gap.

[Carstensen et al. \(2011\)](#) compared the ability of indicators to forecast industrial production in the Euro area. The authors were unable to determine any indicator that was dominant as the best predictor of the industrial production because it depends on the forecast horizon and the loss function considered. Additionally, the forecast of the AR(1) model is quite difficult to beat during quiet times based on the fluctuation test by [Giacomini and Rossi \(2010\)](#). [Rossi and Sekhposyan \(2010\)](#) found that the useful predictors for forecasting US industrial production change over time. However, they did not use a disaggregated series of industrial production as [Carstensen et al. \(2011\)](#) did. [Kotchoni et al. \(2019\)](#) analyzed the performance of models selecting factors from 134 monthly macroeconomic and financial indicators to forecast industrial production, and they compared these models to standard time series models. They found that the MCS selected the LASSO model for forecasting during periods of recessions, but did not choose it to forecast the full out-of-sample data.

When addressing the forecast ability of other economic variables, [Marcellino et al. \(2003\)](#) found evidence that the individual estimation of inflation in each Euro area country and the subsequent aggregation of projections increase the forecast accuracy related to forecasting of this variable at the aggregate level. [Hubrich \(2005\)](#) determined that aggregating the forecasts of each component of inflation does not necessarily better predict inflation in the Euro area one year ahead. [Espasa et al. \(2002\)](#) had similar results, indicating that disaggregation leads to better projections for periods longer than one month. [Carlo and Marçal \(2016\)](#) compared forecasts from models for aggregate inflation and those aggregating the forecasts for the components from the Brazilian inflation index. The authors determined that the forecast using disaggregated data increased accuracy, such as [Heinisch and Scheufele \(2018\)](#).

[Zellner and Tobias \(2000\)](#) studied the effects of aggregated and disaggregated models in forecasting the average annual GDP growth rate of 18 countries. In general, disaggregation led to more observations that could be used to estimate the parameters, but the authors obtained better predictions for the aggregate variable. [Barhoumi et al. \(2010\)](#) analyzed the forecasting performance of France's GDP between alternative factor models. They wanted to know whether it was more appropriate to extract factors from aggregate or disaggregated data for forecasting purposes. Rather than using 140 disaggregated series, [Barhoumi et al. \(2010\)](#) showed that the static approach of [Stock and Watson \(2002\)](#) using 20 aggregate series led to better prediction results. In other words, the mentioned articles present favorable evidence for the use of a disaggregated series or to model using an aggregated series only, leaving the question open.

[Hendry and Hubrich \(2011\)](#) proposed an alternative use of a disaggregate variable to forecast the aggregate variable, which was a combination of disaggregated and aggregated variables. This is different from previous literature, which suggested forecasting the disaggregate variables and then aggregating them to obtain the forecast of the aggregate variable, as researchers discussed earlier in this section. [Hendry and Hubrich \(2011\)](#) determined that including disaggregate variables in the aggregate model improves the forecast accuracy if the disaggregates have different stochastic structures and if the components are interdependent, according to Monte Carlo simulations. They sought to forecast US inflation by considering the sectorial breakdown of inflation. To reduce the dimension of the disaggregate variables, they used the factor model with the results of using this combination, corroborating those obtained by the Monte Carlo simulations. [Hendry and Hubrich \(2011\)](#) introduced (as a promising direction for procedures) selection of the disaggregated series and their lags together with the lags of the aggregate series to predict the aggregate series.

[Faust and Wright \(2013\)](#) analyzed the forecasting models for the US CPI. They considered the combination idea from [Hendry and Hubrich \(2011\)](#) and compared the use of the aggregated or disaggregated series individually in the model, but did not suggest procedures for variable selection. They determined that the combination model did not lead to a better forecasting performance for the aggregated series according to the root of the MSE when compared to disaggregated or aggregated models. [Weber and Zika \(2016\)](#) sought to forecast general employment in Germany as a function of its lags and disaggregation in different sectors. However, the authors used principal components to summarize information from the sectors. They determined that the disaggregation improved the forecast for general employment when compared to the univariate model for the aggregate series. As such, the contributions of this entry include the results of combining the aggregated and disaggregated series and using the variable selection procedure to fill this gap.

Regarding the literature on the methodologies used in this work, [Epprecht et al. \(2021\)](#) conducted a Monte Carlo simulation experiment that considered the data generating process (DGP) to be a linear regression with orthogonal variables and independent data. The authors determined that adaLASSO and the Autometrics algorithms also have similar forecasting performances when there are a small number of relevant variables and when the number of candidate variables is lower than the number of observations. The Autometrics algorithm only performs better when it has a large number of relevant variables (as 15 to 20) because of the bias against the penalization term in adaLASSO. Additionally, [Epprecht et al. \(2021\)](#) determined that adaLASSO performs better than LASSO and the

Autometrics algorithm for linear regression with orthogonal variables in terms of model performance. Autometrics is only preferable with small samples. The authors also used genomics data to compare the predictive power to the epidermal thickness in psoriatic patients, in which covariates are not orthogonal. Out-of-sample forecasts with variables that were selected via LASSO, adaLASSO, or Autometrics cannot be statistically differentiated by the MDM test.

[Kock and Teräsvirta \(2014\)](#) used a neural network model with three algorithms to model monthly industrial production and unemployment series from the Group of Seven (G7) countries and Denmark, Finland, Norway, and Sweden. They focused on forecasting during the economic crisis from 2007 to 2009. The authors found that the Autometrics algorithm performs worse with direct forecasts than with recursive forecasts because the model is not a reasonable approximation of reality (as it excludes the most relevant lags).¹ The Autometrics algorithm tends to select a highly parameterized model that does not present competitive forecasts compared to other methodologies in direct forecasting. That is, [Kock and Teräsvirta \(2014\)](#) determined that the Autometrics algorithm may perform worse when there are considerable misspecifications in the general model. In the present work, researchers used recursive forecasting, in which, according to [Kock and Teräsvirta \(2014\)](#), the Autometrics algorithm does not perform badly.

¹ Direct forecasts require estimating a separate time series model for each forecasting horizon; the only change between each model is the number of horizons ahead for the dependent variable. The recursive forecast is defined if researchers re-estimate the model for each period in the forecast evaluation sample and if researchers compute forecasts with the recursively estimated parameters. See pages 30 and 31 of [Ghysels and Marcellino \(2018\)](#) for a definition of a recursive forecast.

References

1. Lütkepohl, Helmut. 1987. *Forecasting Aggregated Vector ARMA Processes*. New York: Springer Science & Business Media, vol. 284.
2. Lütkepohl, Helmut. 1984. Linear transformations of vector arma processes. *Journal of Econometrics* 26: 283–93.
3. Granger, Clive W. J. 1987. Implications of aggregation with common factors. *Econometric Theory* 3: 208–22.
4. Pesaran, M. Hashem, Richard G. Pierse, and Mohan S. Kumar. 1989. Econometric analysis of aggregation in the context of linear prediction models. *Econometrica: Journal of the Econometric Society* 57: 861–88.
5. Van Garderen, Kees Jan, Kevin Lee, and M. Hashem Pesaran. 2000. Cross-sectional aggregation of non-linear models. *Journal of Econometrics* 95: 285–331.

6. Giacomini, Raffaella, and Clive W. J. Granger. 2004. Aggregation of space-time processes. *Journal of Econometrics* 118: 7–26.
7. Giacomini, Raffaella. 2015. Economic theory and forecasting: Lessons from the literature. *The Econometrics Journal* 18: C22–C41.
8. Hendry, David F., and Kirstin Hubrich. 2011. Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. *Journal of Business & Economic Statistics* 29: 216–27.
9. Aguiar, Mark, and Gita Gopinath. 2007. Emerging market business cycles: The cycle is the trend. *Journal of Political Economy* 115: 69–102.
10. Kohn, David, Fernando Leibovici, and Håkon Tretvoll. 2021. Trade in commodities and business cycle volatility. *American Economic Journal: Macroeconomics* 13: 173–208.
11. Bulligan, Guido, Roberto Golinelli, and Giuseppe Parigi. 2010. Forecasting monthly industrial production in real-time: From single equations to factor-based models. *Empirical Economics* 39: 303–36.
12. Espasa, Antoni, Eva Senra, and Rebeca Albacete. 2002. Forecasting inflation in the european monetary union: A disaggregated approach by countries and by sectors. *The European Journal of Finance* 8: 402–21.
13. Marcellino, Massimiliano, James H. Stock, and Mark W. Watson. 2003. Macroeconomic forecasting in the Euro area: Country specific versus area-wide information. *European Economic Review* 47: 1–18.
14. Hubrich, Kirstin. 2005. Forecasting Euro area inflation: Does aggregating forecasts by hicp component improve forecast accuracy? *International Journal of Forecasting* 21: 119–36.
15. Carlo, Thiago Carlomagno, and Emerson Fernandes Marçal. 2016. Forecasting Brazilian inflation by its aggregate and disaggregated data: A test of predictive power by forecast horizon. *Applied Economics* 48: 4846–60.
16. Heinisch, Katja, and Rolf Scheufele. 2018. Bottom-up or direct? forecasting German GDP in a data-rich environment. *Empirical Economics* 54: 705–45.
17. Quaedvlieg, Rogier. 2021. Multi-horizon forecast comparison. *Journal of Business & Economic Statistics* 39: 40–53.
18. Barnett, Alina, Haroon Mumtaz, and Konstantinos Theodoridis. 2014. Forecasting UK GDP growth and inflation under structural change. a comparison of models with time-varying parameters. *International Journal of Forecasting* 30: 129–43.
19. Hyndman, Rob J., Anne B. Koehler, Ralph D. Snyder, and Simone Grose. 2002. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal*

of forecasting 18: 439–54.

20. Hyndman, Rob, Anne B. Koehler, J. Keith Ord, and Ralph D. Snyder. 2008. *Forecasting with Exponential Smoothing: The State Space Approach*. New York: Springer Science & Business Media.
21. Hyndman, Robin John, and Yeasmin Khandakar. 2008. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software* 27: 1–22.
22. Diebold, Francis, and Roberto Mariano. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13: 253–63.
23. Hansen, Peter R., Asger Lunde, and James M. Nason. 2011. The model confidence set. *Econometrica* 79: 453–97.
24. Harvey, David I., Stephen J. Leybourne, and Paul Newbold. 1998. Tests for forecast encompassing. *Journal of Business & Economic Statistics* 16: 254–59.
25. Faust, Jon, and Jonathan H. Wright. 2013. Forecasting inflation. In *Handbook of Economic Forecasting*. Amsterdam: Elsevier, vol. 2, pp. 2–56.
26. Weber, Enzo, and Gerd Zika. 2016. Labour market forecasting in Germany: Is disaggregation useful? *Applied Economics* 48: 2183–98.
27. Elliott, Graham, and Allan Timmermann. 2008. Economic forecasting. *Journal of Economic Literature* 46: 3–56.
28. Carstensen, Kai, Klaus Wohlrabe, and Christina Ziegler. 2011. Predictive ability of business cycle indicators under test: A case study for the Euro area industrial production. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)* 231: 82–106.
29. Giacomini, Raffaella, and Barbara Rossi. 2010. Forecast comparisons in unstable environments. *Journal of Applied Econometrics* 25: 595–620.
30. Rossi, Barbara, and Tatevik Sekhposyan. 2010. Have economic models' forecasting performance for us output growth and inflation changed over time, and when? *International Journal of Forecasting* 26: 808–35.
31. Kotchoni, Rachidi, Maxime Leroux, and Dalibor Stevanovic. 2019. Macroeconomic forecast accuracy in a data-rich environment. *Journal of Applied Econometrics* 34: 1050–72.
32. Zellner, Arnold, and Justin Tobias. 2000. A note on aggregation, disaggregation and forecasting performance. *Journal of Forecasting* 19: 457–65.
33. Barhoumi, Karim, Olivier Darné, and Laurent Ferrara. 2010. Are disaggregate data useful for factor analysis in forecasting French GDP? *Journal of Forecasting* 29: 132–44.

34. Stock, James H., and Mark W. Watson. 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics* 20: 147–62.
35. Epprecht, Camila, Dominique Guegan, Álvaro Veiga, and Joel Correa da Rosa. 2021. Variable selection and forecasting via automated methods for linear models: Lasso/adalasso and autometrics. *Communications in Statistics-Simulation and Computation* 50: 103–22.
36. Kock, Anders Bredahl, and Timo Teräsvirta. 2014. Forecasting performances of three automated modelling techniques during the economic crisis 2007–2009. *International Journal of Forecasting* 30: 616–31.

Retrieved from <https://encyclopedia.pub/entry/history/show/61966>