

Toppur, Badri; Thomas, T. C.; González-Tejero, Cristina Blanco; Gavrila Gavrila, Sorin

Article

Forecasting commercial vehicle production using quantitative techniques

Contemporary Economics

Provided in Cooperation with:

University of Finance and Management, Warsaw

Suggested Citation: Toppur, Badri; Thomas, T. C.; González-Tejero, Cristina Blanco; Gavrila Gavrila, Sorin (2023) : Forecasting commercial vehicle production using quantitative techniques, Contemporary Economics, ISSN 2300-8814, University of Finance and Management in Warsaw, Faculty of Management and Finance, Warsaw, Vol. 17, Iss. 1, pp. 10-23, <https://doi.org/10.5709/ce.1897-9254.496>

This Version is available at:

<https://hdl.handle.net/10419/297618>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Primary submission: 21.01.2022 | Final acceptance: 02.02.2023

Forecasting Commercial Vehicle Production Using Quantitative Techniques

Badri Toppur¹ and T. C. Thomas¹

Guest editors: Cristina Blanco González-Tejero² and Sorin Gavrila²

ABSTRACT

Firms selling commercial vehicles often face difficulties due to recessions in the globalized economy. Manufacturers are keen to anticipate demand in future quarters to optimize their production schedules. In this study, commercial vehicle production data from a leading Indian automotive manufacturer were analyzed using moving averages, exponential smoothing, seasonal decomposition and autoregressive integrated moving average (ARIMA) models with the goal of forecasting. The results reveal that the ARIMA (0,1,1) model effectively predicts the sectoral downturn coinciding with the global financial crisis of 2008. As life returns to normal after the financial crisis caused by COVID-19, such models may be used to strategically move past the disruption.

KEY WORDS:

demand forecasting, Python programming language, seasonal decomposition, Box-Jenkins methodology.

JEL Classification: C22, B23, G01

¹Rajalakshmi School of Business, Chennai, India

²Economics and Business Management Department, University of Alcalá, Madrid, Spain

1. Introduction

The present study is carried out using a dataset made available by a leading automotive company with advanced engineering and manufacturing capabilities and a local source network. The company designs, manufactures and markets conventional braking products, advanced braking systems, and related air-assisted products and systems, directly serving customers locally and internationally. The company has a major share of the original equipment manufacturer (OEM) market in India as well as the domestic after-sales market, which it serves through a strong and very wide distribution network. Almost 75% of the company's sales are from original equipment manufacturers. Any uncertainty in commercial vehicle production has a significant effect on company sales, as the company depends heavily on this segment. Management is interested in forecasting vehicle sales and controlling production using a study of the trends in the Indian

economy, the commercial vehicle industry scenario, a field survey with fleet operators to gauge market scenarios, a vehicle production forecast, and an original equipment sales forecast.

The company collected data from various OEMs about factors affecting the industry, such as commercial vehicle production, freight rate, diesel prices, emission norms, and the prime lending rate (PLR). In this study, we used 195 observations from approximately sixteen years of records. Table 1 displays a few of the rows of the dataset. Figure 1 shows the time series plot of the commercial vehicle production over a period of 195 months. It is observed that there is a dip around month no. 165, which is uncharacteristic of the overall time series but corresponds in time to the global finance crisis of 2006-08. Although the global crisis was reported as being primarily concerned with housing finance in the USA, it appears to have impacted the banking sector and many other sectors in other countries. In this study, we used traditional techniques to

Correspondence concerning this article should be addressed to:

Badri Toppur, Rajalakshmi School of Business, Kuthambakkam Post, NH-4, Chembarambakkam, Chennai 600124, India.

E-mail: badri.toppur@rsb.edu.in

forecast commercial vehicle sales over that historical period. Figure 2 shows that there is a significant correlation between the original series and lagged values of the series. This correlation indicates the suitability of applying time series methodology to the problem. The Box-Jenkins methodology was used to specify and fit ARIMA models.

The study concludes that, based on common forecasting error measures such as the mean absolute percentage error (MAPE) and mean squared error (MSE), the double exponential smoothing method forecasts the series in the disrupted period better than any of the moving average methods. The Box-Jenkins method, although more complex than simple moving averages and double exponential smoothing, predicts the downturn in the company's revenue well using various criteria for comparing models such as the Akaike information criterion (AIC) and the Schwarz Bayesian information criterion (BIC). The fact that the downturn in the vehicle company's time series coincided with the housing finance crisis, whereas the price of diesel remained unaffected, suggests that many commodities in the category of high-priced consumer durables

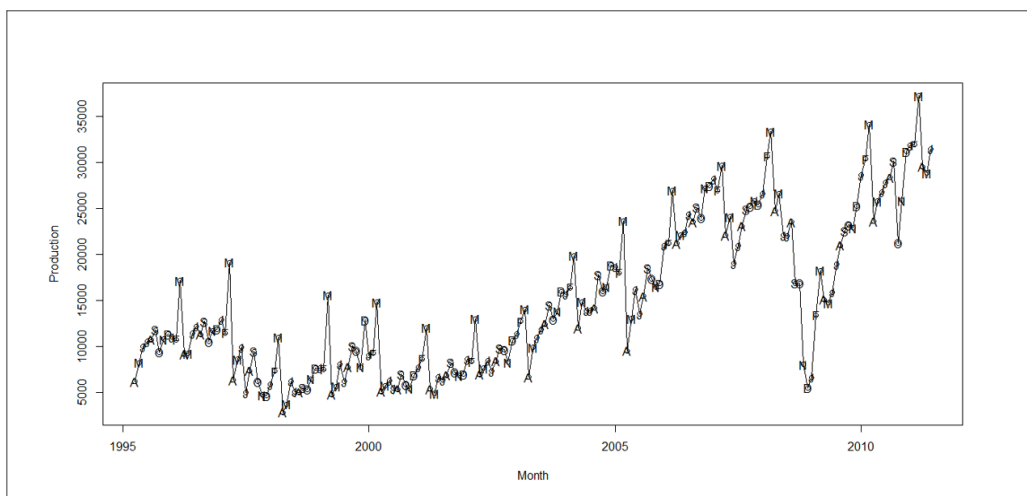
are affected simultaneously in a global financial crisis, whereas demand for fast moving consumer goods (FMCG) is relatively inelastic.

This paper is organized as follows. Section 1 presents the introduction to the problem, followed by the literature review in Section 2. Section 3 explains the methodology used; Sections 4 and 5 present the results and their discussion, respectively. Conclusions and recommendations are presented in Section 6, followed by the limitations of the study in Section 7.

2. Literature Review

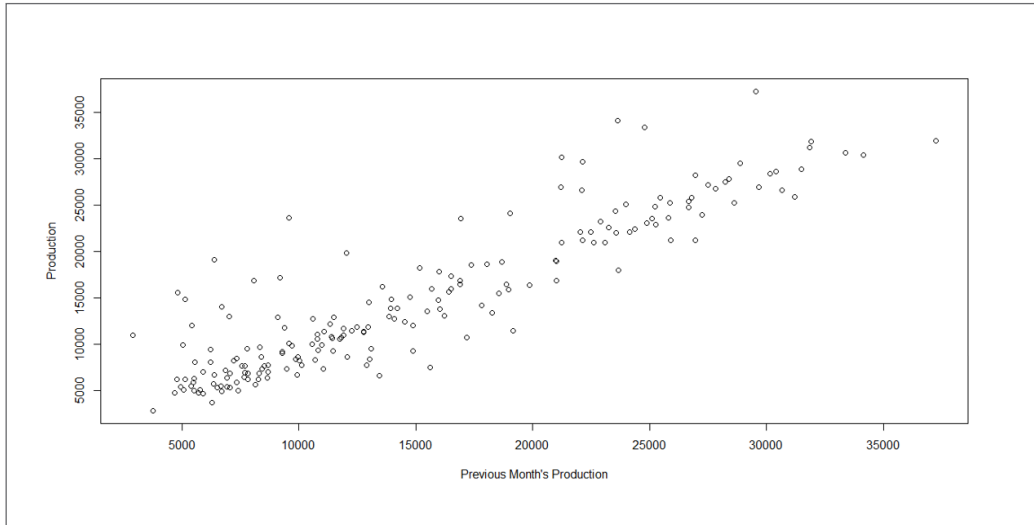
There have been positive and negative consequences of globalization, and it is a subject of ongoing debate among economists. Globalization has created employment opportunities worldwide, which is a plus. The insistence on unbalanced free trade agreements has also created some problems for third-world countries. Stiglitz (2012) puts market economics and the globalization phenomenon in an accurate perspective for developing countries and discusses the financial crisis in East Asia. Wilczyński (2011) presents another view of the International Monetary

Figure 1
Time Series of Length 195



Source: R package TSA

Figure 2
Scatterplot of Production vs. First Lag of Production



Source: R package TSA

Table 1
Head of the Dataset

Period	Production	Index of Industrial Production	Diesel Price	Freight Index	Prime Lending Rate
1	6224	115.3	6.99	100.00	16.50%
2	8262	114.9	7.02	100.50	16.50%
3	10003	113.0	7.13	101.25	16.50%
4	10567	116.6	7.08	107.81	16.25%
5	10793	116.7	7.20	113.45	16.25%
6	11811	120.0	7.28	124.59	16.25%
7	9393	116.9	7.32	121.88	16.00%
8	10811	123.1	7.37	153.85	15.50%
9	11382	132.5	7.40	122.09	15.00%
10	11061	135.1	7.38	127.03	15.00%
11	10785	131.8	7.40	122.09	15.00%
12	17176	144.2	7.41	121.46	15.00%
13	9195	128.4	7.43	113.84	15.00%
14	9295	129.8	7.46	112.90	15.00%
15	11459	124.3	7.49	112.06	15.00%

Fund (IMF) that was criticized by Stiglitz for its ideological implementation of policies. Globalization has created many interdependencies between the banking systems of different countries. This sectoral study is from the times of the global financial crisis (GFC) of 2007-09. The banking crash caused by the aggressive housing finance in the USA had repercussions in the stock markets and on many foreign banks that were exposed to it. Through the banking system, the GFC affected many industrial sectors, including manufacturing. It is considered to have been as severe as the Great Depression of the 1930s.

The impact of the crash in a few regions outside the US can be visualized through the bank reports of France and Australia, as presented below.

As described by the Banque de France (2008), "Falling house prices in the USA from 2006, coupled with a gradual rise in interest rates, caused large numbers of borrowers to default on their loans, which in turn resulted in failures by mortgage lenders. The major US banks that either financed these institutions or were shareholders in them also suffered losses and asset impairment. Consequently, the problems on the mortgage lending market swiftly turned into a global financial and stock market crisis. The crisis spread to other regions of the world and other banking systems because of the massive securitization of subprime loans, which became extremely popular owing to their high rates of return."

According to the Reserve Bank of Australia (2012), "Australia did not experience a large economic downturn or a financial crisis during the GFC. However, the pace of economic growth did slow significantly, the unemployment rate rose sharply, and there was a period of heightened uncertainty. The relatively strong performance of the Australian economy and financial system during the GFC, compared with other countries, reflected a range of factors, including the fact that Australian banks had very small exposures to the US housing market and US banks, partly because domestic lending was very profitable. Subprime and other high-risk loans were only a small share of lending in Australia, partly because of the historical focus on lending standards by the Australian banking regulator (the Australian Prudential Regulation Authority [APRA]). Australia's economy was buoyed by large resource exports to China, whose economy re-

bounded quickly after the initial GFC shock (mainly due to expansionary fiscal policy)."

These extracts show that the GFC did not greatly affect all countries, but only those that had considerable financial institution linkages to the US. The securitization mentioned in the first quotation occurred through mortgage-backed securities (MBSs) that were sold in the US and many European countries. In truth, many of the mortgages were extremely risky, and some were valued at more than the price of the homes.

Vali's (2014) book on mathematical economics stimulates interest in the analysis of mortgage financing. The book's preface notes that "A 1993 Nobel Prize in Economics was awarded to Robert Fogel of the University of Chicago and Douglass North of Washington University, both economic historians, for their contribution to the application of quantitative methods to the field of economic history. Fogel and North advanced 'Cliometrics', a new quantitative technique for analyzing historical data and solving historical puzzles." A chapter in this textbook discusses refinancing options and equal monthly installments (EMIs) for long-term loans or mortgages. It demonstrates that a new plan offering lower interest rates may not always be beneficial and that there are some conditions—regarding the principal and interest that one has already paid—that must be met if refinancing the remaining part of the loan or mortgage is to truly benefit the investor. Thus, a well-intentioned economic practice such as lending to economically weaker sections (EWSs) may cause a global crisis if the bank is not careful. The coronavirus pandemic has had a tremendous impact on the world economy, and although evidently a health crisis, this impact suggests that there are some similar compelling and widespread practices, perhaps not directly economic, that have precipitated it. Skare and Stjepanović (2016) have discussed business cycles with the aim of identifying the determinants. They note that although frameworks exist for measuring business cycles, none are perfect. Frequency domain analysis is an option for detecting cycles, for example, in world GDP. Historically, classical economists have denied business cycles outside of a war situation. Currently, business cycles are considered to be caused by technological shocks, shocks specifically in the oil industry, or financial or mon-

etary shocks in general. In another interesting article, Makovský (2018) examines Mankiw's puzzle about the growth of durable consumption during a recession. This study examines the scrapping of old cars and new car registration data from recession periods in a few European economies. Sinković et al. (2021) monitored economic indicators in the context of financial and economic crises. They looked specifically at the federal interest rates in the US before, during and after the Great Depression. They also studied the changes in economic variables such as GDP, inflation, consumer prices, federal debt and the S&P 500 index in the period of the great recession of 2007–2009 and the recent COVID-19 crisis.

Anderson (1977) has explained the basic difference between time series analysts and other statisticians. Time series analysts are interested in the correlation between values at separate times to forecast future values of the series. In probability and statistics, however, one often assumes or proves the independence of the random variables, such as in regression techniques. Another paper explains the basics of the Box-Jenkins methodology, dwelling on the difficult choice of parameter values for the various models. Here, the ramifications of the $p + q \leq 2$ inequality are explored, where p and q are the parameters of the autoregressive moving average (ARMA) models. In the literature, one can find models that have $p = q = 6$, for example, in Viccione et al. (2020). To determine the best values for p and q , measures such as AIC, BIC, and the Hannan-Quin information criterion (HQIC) are used. Johansen (2011) also warns against the use of regression and correlation methods with nonstationary data and compares them unfavorably to cointegrated vector autoregression models. Differencing the time series is a method used to address stationary time series, as explained by Pradhan and Kumar (2021). P. Selva (2019) wrote a tutorial on time series analysis using Python. James et al. (2013) describe the basics of data analysis in the text *Elements of Statistical Learning*. Their discussion of flexible and inflexible models with respect to degrees of freedom and the effect of that decision on training and test mean square error (MSE) is quite illuminating. Render et al. (2009) explain the various forecast error formulae. Time series analysis using the R language is described in the revised edition of Cryer and Kung-Sik Chan

(2008) with the associate R package for time series analysis called TSA. Jin Lian and Ling He (2018) document an optimization method for selecting the α parameter for exponential smoothing. The state of the art in forecasting can be found in the paper by T. M. Dantas and F. L. Cyrino Oliveira (2018), which details an approach combining bootstrap aggregation, clusters, and exponential smoothing. Byeongchan Seong (2020) recently developed vector exponential smoothing. Slawek Smyl (2020) created a hybrid method of exponential smoothing and recurrent neural networks.

The literature reviewed thus far explains the theoretical concepts of the various statistical approaches used for forecasting, whereas the following briefly presents their specific applications.

Ampountolas (2021) used SARIMAX, neural networks, and GARCH models to forecast daily hotel demand. Benvenuto et al. (2020) applied the ARIMA model to COVID-19 epidemic data. Chen and Yu (2012) examined the global financial crisis of 2008. Chodakowska et al. (2021) used ARIMA models in electrical load forecasting. Duan et al. (2022) attempted to predict gas well production using ARIMA-RTS models. Fan et al. (2021) used ARIMA-LSTM models, which also performed well in production forecasting. Fattah et al. (2018) forecast the demand for food products. Giri et al. (2022) developed a deep learning model for demand forecasting in the fashion and retail industries. Gu (2022) examined the impact of the sharing economy on urban carbon emissions from 2011 to 2017 using a sample of 275 prefecture-level cities in China. Haque and Shaik (2021) predicted crude oil prices during the epidemic using ARIMA and GARCH models. Irhami and Farizal (2021) forecast the number of vehicles in Indonesia using the ARIMA method. Khan and Alghulaiakh (2020) used ARIMA for forecasting stock prices. Lee and Suh (2022) suggested a method for modeling causality between environment, social and governance (ESG) conduct and financial performance variables. Lewicki and Olejarz-Wahba (2020) modeled the trend in sales of new passenger vehicles in Poland. Li and Li (2017) forecast energy consumption in Shandong, China. Menculini et al. (2021) compared Prophet software from Facebook and deep learning methods to ARIMA for forecasting wholesale food prices. Pe-

dregal (2020) used multiple forecasting methods for predicting uranium prices. Nafil et al. (2020) forecast energy demand in Morocco. Suraiya and Hasan (2018) predicted the requirement of printer paper in universities of Bangladesh based on 6 years or 18 trimesters of data. Sidqi and Sumitra (2019) used exponential smoothing methods on twenty-four observations corresponding to weeks of sales data. Taques, et al. (2021) examined the phenomenon of innovation. Innovation can be a source of competitive advantage for companies, either through the improvement of methods and techniques capable of generating new products or services or through perfecting existing ones. Temür et al. (2019) predicted housing sales in Turkey using ARIMA, LSTM, and hybrid models. Trapp et al. (2022) combined business model theory with the concept of sector coupling in their study, which identifies 12 business model archetypes in the emerging hydrogen vehicle industry and its value chain. Vo et al. (2021) applied the Box-Jenkins model and fuzzy EPQ model to the apparel industry. Zeng et al. (2019) benchmarked an automotive company's performance from economic and environmental perspectives by creating a relevant index.

3. Methodology

One intriguing theory in forecasting methodology is that the effect of the many possible economic variables impacting the response variable is captured in the auto-correlations within the time series. One does not need to model the response variable explicitly as a function of many variables. Even without all these independent variables, the selection of a suitable model among the many available models remains a difficult iterative procedure. The number of parameters to be estimated can influence the bias and variance, and thus, the flexibility of the model needs to be decided after analyzing the data. Subsequent to the choice of model, forecasting can be performed and bounds on the error can be estimated.

3.1. Seasonal Decomposition

These time series are usually understood to be the composition of a trend component (T), a seasonality component (S), an aperiodic cycle component (C), and a random noise component (R). These components may be additive, as $T + S + C + R$, but a more realistic model is their product form $T \times S \times C \times R$. In the *statsmodels*

package in Python, there are functions that easily decompose a loaded time series into the constituent components, which can then be plotted independently to see the pattern in the component

3.2. Moving Average and Exponential Smoothing

The simpler methods, such as moving averages and exponential smoothing, use the simple concept of a weighted average. The weights attached to various values in the weighted average can be equal, or one may favor more recent events with greater weight. In exponential smoothing, there are one or more parameters used, with which one can estimate the trend and the next period forecast. In simplex exponential smoothing, there is only one smoothing constant α that is between 0 and 1. One takes an alpha portion of the forecast error in the previous period and adds it to the period's forecast value to obtain the current period forecast. The alpha parameter is chosen to minimize the mean absolute deviation (MAD), the formula for which is given in Section 3.1.1. More complex variations of simple exponential smoothing include Holt's method and Winter's method. Holt's method refers to trend-adjusted smoothing, which is also called second-order smoothing or double smoothing. A second parameter, β , is used to adjust the trend term in the forecast. Instead of simply updating the forecast term F_t , one also separately updates the trend term T_t as a fraction of the error in the trend term. The acronym FIT_t stands for forecast including trend and is the sum of F_t and T_t . Higher values of beta give more importance to recent changes in trend, whereas low values of beta smooth the changes in trend. Again, the parameter values are chosen to minimize MAD. Winter's method is built on top of Holt's method; therefore, the method is also referred to as Holt-Winter smoothing or triple exponential smoothing. The added aspect of seasonality is also addressed in Winter's method.

3.2.1. Forecast error measures

In time series terminology, bias is simply the average of the error terms F_t and Y_t . In all the formulae given in this section, Y_t is the actual series value at time t and F_t is the forecasted series value at time t . The formula for the mean absolute deviation is as follows:

$$MAD = \frac{\sum |Y_t - F_t|}{n}$$

In addition, two common error measures used are the mean squared error (MSE) and the mean absolute percentage error (MAPE).

$$MSE = \frac{(F_t - Y_t)^2}{n}$$

$$MAPE = \frac{\sum \left| \frac{F_t - Y_t}{Y_t} \right|}{n} 100\%$$

3.3. Box-Jenkins Models

Definition: A sequence of random variables $\{Y_t, t = 0, \pm 1, \pm 2, \pm 3, \dots\}$ is called a stochastic process and is representative of a time series.

Definition: The autocorrelation function is $\rho_{t,s} = \text{Corr}(Y_t, Y_s)$ for $t, s = 0, \pm 1, \pm 2, \pm 3, \dots$

Stationarity is an important assumption in these models. In layperson's terms, stationarity means that the probabilistic behavior of the time series is independent of the time window in which it is being observed if the length of the window is not changed. If this property is present, it is more conducive to model building and forecasting. The acronyms ACF and PACF are expanded as sample autocorrelation function and sample partial autocorrelation function, respectively. These correlograms, as they are otherwise referred to, indicate the correlation between lagged values of the time series. Along with the augmented Dickey-Fuller test, they help one decide about the stationarity of the series and estimate the p and q parameters of the AR and MA components of the ARIMA model. The AIC and BIC are used to compare the ARIMA models with differing parameters (p, d, q) by penalizing overfitting. In the following formulae, k is a simple linear function of p and q . The maximum likelihood estimate is the difference between the observed and expected distribution of the stochastic process. The calculation of the maximum likelihood estimate is mathematically quite complex compared to ordinary least squares estimates and is not described here.

$$AIC = -2\text{Log}(\text{Maximum Likelihood}) + 2k$$

$$BIC = -2\text{Log}(\text{Maximum Likelihood}) + k\text{log}(n)$$

Definition: A process $\{Y_t\}$ is said to be strictly stationary if the joint distribution of $Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_n}$ is

the same as the joint distribution of $Y_{t_1-k}, Y_{t_2-k}, Y_{t_3-k}, \dots, Y_{t_n-k}$ for all choices of $t_1, t_2, t_3, \dots, t_n$ and for any lag k .

Definition: A process $\{Y_t\}$ is said to be weakly stationary if two conditions are met:

- The mean function $\mu_t = E[Y_t]$ is constant over time
- $\gamma_{t,t-k} = \gamma_{0,k}$ for all time t , and lag k

In $\gamma_{t,s}$ is the covariance function, which is related through the variances to the correlation function $\rho_{t,s}$. If multivariate normality assumptions are not met or have not been tested, the weaker definition of stationarity is implied.

The Box-Jenkins methodology entails model specification, model fitting, and model diagnostics, which may be repeated until a good model is identified. The specification requires a selection among the various forecasting models that are available, such as pure moving averages, pure autoregression models, hybrid ARMA models, and integrated ARIMA models. Model fitting requires that the parameters of the model be estimated accurately, and diagnostics validate the assumptions of the model and its forecasting accuracy.

The Python functions for ARIMA are given in the output measure for testing model assumptions, such as the Ljung-Box test for white noise or independence of the random variables comprising the series. The alternate hypothesis is that the serial data are correlated. There is also a heteroscedasticity test for constant variance and the Jarque-Bera test for normality of errors or residuals. If the test statistic in the Jarque-Bera test is large, then one concludes that the residuals are not normally distributed. These output measures considered together can help one conclude whether the model chosen is robust.

4. Results

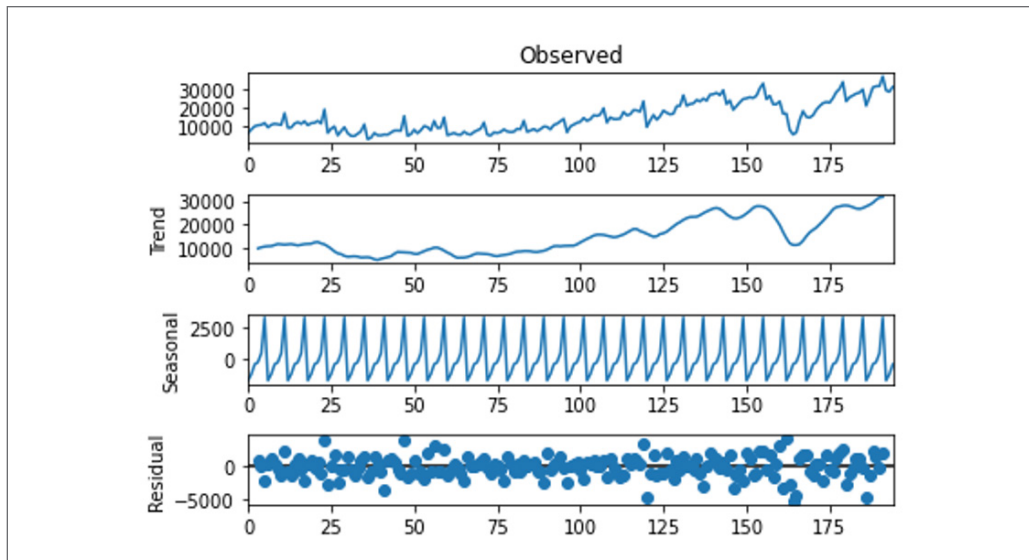
In this paper, we have applied the primary forecasting methods in the Python package *statsmodels.api*, particularly the *tsa* package. These are the moving averages, exponential smoothing, and the numerous ARIMA models that may be used exclusively as an autoregressive model, a moving average model, or an integrated model. It was found through experimental trials that the additive model was a better fit than the multiplicative model, as indicated by the residual pattern, which is ran-

domly distributed about the x-axis. In the seasonal decomposition in Figure 3, the trend line, which is fairly level, begins climbing from the 75th month to approximately the 160th month. Then, the trend line has a convex-upward shape or minima at ap-

proximately the 165th month before it picks up and returns to climbing as before. The seasonality is remarkably consistent. An additive model instead of a multiplicative model yielded residuals that were random.

Figure 3

Seasonal Decomposition - Additive Model



Source: Python package statsmodels.tsa.seasonal

The moving average gave an MAPE of 8.71% to 13.43, with the lowest value for a span of 7. However, the forecast missed the dip in the time series by quite a few months. Nor did simple exponential smoothing perform well in predicting the months in question. The errors decrease monotonically with increasing α . Table 2 shows the sensitivity analysis for double exponential smoothing using the Solver Table from Palisade Tools. For different values of the smoothing constants, we can see the corresponding MAD errors. In this table, one can see that the optimal choice of least MAD is for $\alpha = 0.7$ and $\beta = 0.1$. This conclusion is supported by the optimal choice of least MSE (not shown),

which also selects $\alpha = 0.7$ and $\beta = 0.1$.

Moving on, past the exponential smoothing to the Box-Jenkins methodology, the ACF plot of the time series revealed that the first fourteen correlations are outside the lower and upper limits, indicating that the series is not stationary. The PACF plot showed that the first two correlations are outside the lower and upper limits, with the third on the borderline. For this reason, one would suggest setting $p = 2$ in the ARMA model. Empirical studies by Box and Jenkins have shown that the sum of the parameters p and q does not exceed 2 for a good model that is parsimonious and not overfitted.

Table 2
Sensitivity Analysis for Holt's Method

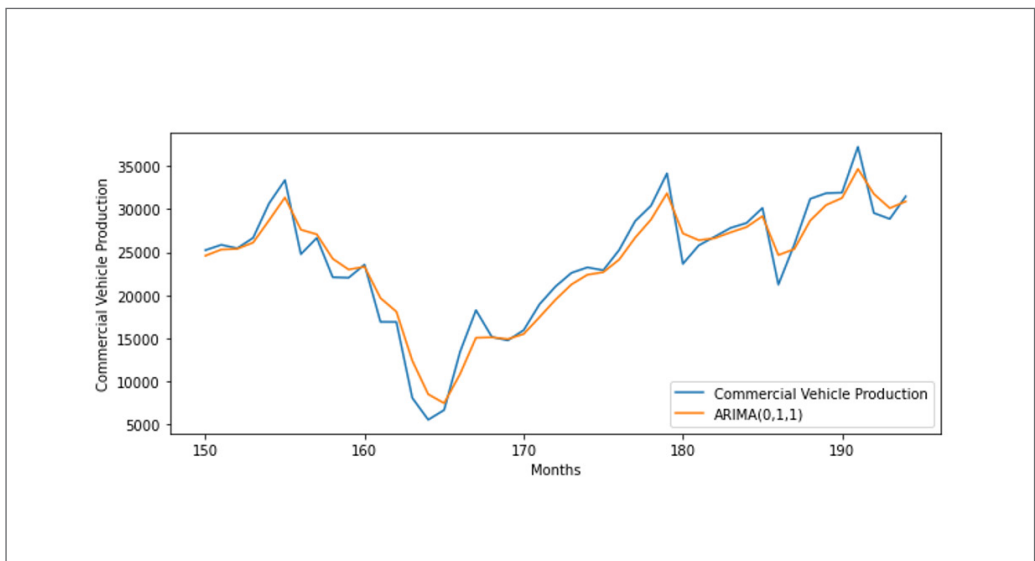
beta	alpha			MAD					
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	3264.16	2770.26	2615.41	2479.71	2368.08	2309.65	2290.88	2310.35	2365.29
0.2	3442.59	2922.28	2691.96	2502.81	2392.97	2343.07	2342.07	2381.76	2465.12
0.3	3466.35	3003.74	2722.00	2517.19	2414.22	2388.81	2422.97	2477.90	2583.63
0.4	3522.14	3048.39	2749.27	2533.78	2459.02	2468.06	2516.45	2588.55	2723.16
0.5	3539.01	3126.36	2761.75	2562.35	2524.66	2569.49	2610.24	2708.23	2866.50
0.6	3629.48	3181.97	2770.64	2607.00	2609.15	2666.89	2711.26	2830.44	3013.62
0.7	3766.72	3203.31	2786.8	2646.86	2706.01	2753.37	2801.83	2951.51	3159.14
0.8	3892.55	3216.48	2811.71	2712.52	2804.56	2828.94	2895.58	3076.37	3305.78
0.9	3964.28	3239.92	2822.42	2797.39	2900.05	2900.5	2992.17	3201.91	3467.29

Source: Solver Table from Palisade

Table 3 shows by the augmented Dickey–Fuller test that the given series is not stationary but that the first-order differenced series is stationary. Table 4 shows the results of the ARIMA testing with various combinations of the parameter values. There are 18 combinations possible, with $p+q \leq 2$, out of which two involving $p = q = 0$ must be rejected. In

this table, one can see that Model #7 has the lowest AIC and BIC and is therefore selected as the best fitting model. Figure 4 shows an overlay of the forecast values from this model for the last 15 time periods on the original series. The forecast correctly captures the downturn in production around February, March, and April of 2008.

Figure 3
Overlay of IMA (1,1) Forecast on the Time Series



Source: Python package matplotlib

Table 3*Augmented Dickey–Fuller Test for Original Series and First-Order Difference Series*

	Test Statistic	P value	Lags used	Number of Observations
Original Series	-0.22	0.94	15	179
Lagged Series	-4.35	0	14	165

Source: Python package pandas

Table 4*Akaike and Bayesian Information Criterion Scores for ARIMA Models with Different Parameters*

Model #	p, d, q	AIC	BIC	Ljung–Box p value	Jarque–Bera p value
1	1, 0, 0	3740.082	3749.901	0.00	0.00
2	2, 0, 0	3719.683	3732.775	0.39	0.00
3	0, 0, 1	3928.091	3937.910	0.00	0.04
4	0, 0, 2	3936.823	3949.915	0.01	0.85
5	1, 1, 0	3699.163	3705.698	0.28	0.00
6	2, 1, 0	3694.596	3704.400	0.92	0.00
7	0, 1, 1	3692.913	3699.449	0.93	0.00
8	0, 1, 2	3694.885	3704.689	0.86	0.00
9	1, 0, 1	3714.476	3727.777	0.95	0.00
10	1, 1, 1	3694.873	3704.676	0.87	0.00
11	2, 1, 1	3696.407	3709.478	0.87	0.00
12	2, 1, 2	3697.320	3713.660	0.78	0.00
13	1, 1, 2	3695.406	3708.477	0.73	0.00
14	0, 1, 0	---	---	---	---
15	2, 0, 1	3716.473	3732.838	0.94	0.00
16	1, 0, 2	3716.628	3732.993	0.94	0.00
17	2, 0, 2	3709.570	3729.208	0.77	0.00
18	0, 0, 0	---	---	---	--

Source: Python package statsmodels.tsa.arima.model

5. Discussion

In the seasonal decomposition approach shown in Figure 3, an additive model instead of a multiplicative model yielded residuals that were random. This outcome is not common; usually, the multiplicative $T \times S \times C \times R$ is considered to provide a better fit. One possible explanation for this exception to the rule is that the influence of a phenomenon external to the industry on business cycle being is independent of the trend and seasonality. Double

exponential smoothing gave consistent estimates for the smoothing parameters with both error measures that were used. In the Box–Jenkins approach, the ACF plot reveals that the first fourteen bars are outside the stationarity region, and therefore, one could try many different values in the interval $\{0,14\}$ for q in the $MA(q)$ model. However, we adhered to a conservative philosophy and only used lower-order models with the intent of not overfitting the model. Overfitted models do not in general

work well on new test datasets. The PACF indicates that the first two spikes are outside the stationarity region, but the AIC and BIC measures indicate that $p = 0$ is better in the AR(p) model than $p = 1$ or $p = 2$. The augmented Dickey–Fuller tests show that first-order differencing is needed and that the differenced series is stationary. The Ljung–Box test statistic of the test is $Q = 0.01$, and the p value of the test is 0.93, which is much larger than 0.05. Thus, we fail to reject the null hypothesis of the test and conclude that the residual values are independent. As the test statistic in the Jarque–Bera test is large (40.78) and the p value is almost zero, one concludes the alternate hypothesis that the residuals are not normally distributed. The normality assumption of residuals allows statisticians to validate the estimation of the parameters mathematically. If residuals follow some other distribution, distribution fitting software may be used to determine the best-fitting distribution. Validation of the parameter estimates may then be possible, although more difficult than it is for the normal distribution case.

6. Conclusions and Recommendations

Naive forecasting methods such as seasonal decomposition, both additive and multiplicative, as well as moving averages and exponential smoothing have been essayed on an industrial dataset from the time of the global financial crisis of 2008. The sympathetic downturn of the commercial vehicle market around the same period has been highlighted in this paper. The more complex Box–Jenkins methodology for combined autoregressive moving average models has yielded good results for this historical dataset. During the COVID-19 years beginning in 2020 and continuing well into 2021, the sales and functioning of several businesses were affected by the lockdowns, with factory shutdowns, construction project halts, and interstate transportation networks that had to deal with a large, unemployed, migrant workforce that had to return home from the worksites. Many of these factories are in fact a part of a global supply chain. These companies will have resumed operations, but the future trend or direction will be uncertain for some time.

As a policy recommendation, IMA (1,1) models similar to the one used in this paper should be ap-

plied to such COVID-19 data series. New theories, variables, and models will be needed, as the coronavirus crisis was novel in its worldwide destruction and the number of fatalities it caused. Another line of research that can be pursued further is to determine the categories of goods that are affected positively or negatively during a crisis. The demand for consumer durables took a hit during the global financial crisis, whereas the demand for fast-moving consumer goods remained unabated. In the COVID-19 crisis, demand for medical supplies obviously peaked worldwide, but what other types of products were affected positively or negatively in that period is a useful research topic to pursue. Governments and businesses may then be prepared with contingency plans for such synergies.

7. Limitations

The most important lesson is that business cycles such as a recession or contraction caused by a housing finance crisis or a pandemic are so rare compared to seasonality that they are difficult to predict, even with many years of monthly data. One may require a century of data to forecast the next business cycle caused by a pandemic. There are such datasets in the scientific literature, such as astronomical data, that span hundreds of years. However, for any given other type of application, specifically in the economic and social sciences, it is unlikely that anyone has access to such a comprehensive database.

Financial time series are known to be different from demand time series by their extraordinary volatility or variance. Autoregressive conditional heteroscedasticity (ARCH) models and their generalized variants are recommended for modeling these time series. Similarly, the COVID-19-affected time series data may be fundamentally different from the Great Depression period data of 1930 or the global financial crisis data of 2006–08. The waves were extremely difficult to predict given existing epidemiological models for influenza, but in this case, one must remember that we are not forecasting an epidemiological time series but COVID-19-affected demand, production or sales data. This should only affect the model specification, and even if a new class of models is designed specifically for modeling them, the ARIMA methodology will still be of good use to forecasters in the next few years.

References

- Ampountolas, A. (2021). Modeling and forecasting daily hotel demand: A comparison based on SARLMAX, neural networks, and GARCH models. *Forecasting*, 3, 580–595. <https://doi.org/10.3390/forecast3030037>
- Anderson, O. D. (1977). The Box-Jenkins approach to time series analysis. *RAIRO-Operations Research*, 11(1), 3–29. <https://doi.org/10.1051/RO/1977110100031>
- Azgun, J. A., San-Jose, L., & Urionabarrenetxea, S. (2013). The impact of financial globalization and financialization on the economy in the current crisis through banking corporate governance. *Contemporary Economics*, 7(3), 79–94. <https://doi.org/10.5709/ce.1897-9254.91>
- Banelienè, R. (2021). Sustainable economic growth in the European Union under COVID-19 conditions. *Contemporary Economics*, 16(1), 121–134. <https://doi.org/10.5709/ce.1897-9254.472>
- Banque De France. (2008). Impact de la crise financière sur les pays du groupe Afrique francophone au FMI et à la Banque mondiale [Impact of the international financial crisis on the countries of the Francophone African Constituency of the International Monetary Fund and the World Bank]. *Research and International Relations Directorate of the Central Bank of West African States (BCEAO)*. https://www.banque-france.fr/sites/default/files/media/2016/11/02/razf_2008.pdf
- Benvenuto, D., Giovanetti, M., Vassallo, L., Angelletti, S., & Ciccoczi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in Brief*, 29, 105340. <https://doi.org/10.1016/j.dib.2020.105340>
- Chen, Y.-C. & Yu, W.-C. (2012). The housing price forecasting and the out-break of the financial crisis: evidence of the ARIMA model. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1610330>
- Chodakowska, E., Nazarko, J., & Nazarko, L. (2021). ARIMA models in electrical load forecasting and their robustness to noise. *Energies*, 14(23), 7952. <https://doi.org/10.3390/en14237952>
- Cryer, J. D. & Chan, K.-S. (2008). *Time series analysis - with applications in R* (2nd ed.). Springer.
- Dantas, T. M. & Oliveira, F. L. C. (2018). Improving time series forecasting: An approach combining bootstrap aggregation, clusters, and exponential smoothing. *International Journal of Forecasting*, 34, 748–761. <https://doi.org/10.1016/j.ijforecast.2018.05.006>
- Duan, Y., Wang, H., Wei, M., Tan, L., & Yue, T. (2022). Application of ARIMA-RTS optimal smoothing algorithm in gas well production prediction. *Petroleum*, 8(2), 270–277. <https://doi.org/10.1016/j.petlm.2021.09.001>
- Fan, D., Sun, H., Yao, J., Zhang, K., Yan, X., & Sun, Z. (2021). Well production forecasting based on ARIMA-LSTM model considering manual operations. *Energy*, 220. <https://doi.org/10.1016/j.energy.2020.119708>
- Fattah, J., Ezzine, L., Aman, Z., Moussami, H. E., & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management*, 10, 1847979018808673. <https://doi.org/10.1177/1847979018808673>
- Giri, C. & Chen, Y. (2022). Deep learning for demand forecasting in the fashion and apparel retail industry. *Forecasting*, 4(2), 565–581. <https://doi.org/10.3390/forecast4020031>
- Gu, J. (2022). Sharing economy, technological innovation and carbon emissions: evidence from Chinese cities. *Journal of Innovation & Knowledge*, 7(3), 100228. <https://doi.org/10.1016/j.jik.2022.100228>
- Haque, M. I. & Shaik, A. R. (2021). Predicting crude oil prices during a pandemic: a comparison of ARIMA and GARCH models. *Montenegrin Journal of Economics*, 17(1), 197–207. <https://doi.org/10.14254/1800-5845/2021.17-1.15>
- Irhani, E. A. & Farizal, F. (2021). Forecasting the number of vehicles in Indonesia using auto regressive integrative moving average (ARIMA) method. *Journal of Physics: Conference Series*, 1845(1). <https://doi.org/10.1088/1742-6596/1845/1/012024>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
- Johansen, S. (2012). The analysis of nonstationary time series using regression, correlation and cointegration. *Contemporary Economics*, 6(2), 40–57. <https://doi.org/10.5709/ce.1897-9254.39>
- Khan, S. & Alghulaiakh, H. (2020). ARIMA model for accurate time series stocks forecasting. *International Journal of Advanced Computer Science and Applications*, 11(7), 524–528. <https://doi.org/10.14569/IJACSA.2020.0110765>
- Lee, M. T., & Suh, I. (2022). Understanding the effects of environment, social, and governance conduct on financial performance: arguments for a process and integrated modelling approach. *Sustainable Technology and Entrepreneurship*, 1(1), 100004. <https://doi.org/10.1016/j.stae.2022.100004>

- Lewicki, W. & Olejarz-Wahba, A. A. (2020). Modeling of trend in sales of new passenger vehicles in Poland using the ARIMA-X-12 and Tramo-seats methods. *Statistika*, 100(3), 333-344.
- Li, S. & Li, R. (2017). Comparison of forecasting energy consumption in Shandong, China using the ARIMA model, GM model, and ARIMA-GM model. *Sustainability*, 9(7), 1181. <https://doi.org/10.3390/su9071181>
- Lian, J. & He, L. (2018). Research on production prediction based on exponential smoothing method. In *9th International Conference on Medicine and Education, of the IEEE Computer Society* (pp. 961-963). IEE. <https://doi.org/10.1109/ITME.2018.00214>
- Makovský, P. (2018). Mankiw's 'puzzle' – is durable consumption declining? *Contemporary Economics*, 13(4), 471-479. <https://doi.org/10.5709/ce.1897-9254.326>
- Menculini, L., Marini, A., Proietti, M., Garinei, A., Bozza, A., Moretti, C., & Marconi, M. (2021). Comparing prophet and deep learning to ARIMA in forecasting wholesale food prices. *Forecasting*, 3(3), 644-662. <https://doi.org/10.3390/forecast3030040>
- Nafil, A., Bouzi, M., Anoune, K., & Ettalabi, N. (2020). Comparative study of forecasting methods for energy demand in Morocco. *Energy Reports*, 6, 523-536. <https://doi.org/10.1016/j.egy.2020.09.030>
- Nguyen, H., Phan Nha, T. T., & Ngo, V. M. (2022). The impact of the COVID-19 pandemic on the performance and risk profile of airline stocks. *Contemporary Economics*, 16(3), 257-275. <https://doi.org/10.5709/ce.1897-9254.481>
- Pedregal, D. J. (2020). Forecasting uranium prices: Some empirical results. *Nuclear Engineering and Technology*, 52, 1334-1339, <https://doi.org/10.1016/j.net.2019.11.028>
- Pradhan, M. & Kumar, D. (2021). *Machine learning using Python*. Wiley Publishers.
- Render, B, Stair, R. M., Hanna, M. E., & Hale, T. S. (2009). *Quantitative analysis for management* (9th ed.). Pearson Education.
- Reserve Bank of Australia. (n.d.) *The global financial crisis*. Reserve Bank of Australia. <https://www.rba.gov.au/education/resources/explainers/the-global-financial-crisis.html>
- Selva, P. (2019, August 22). ARIMA model - complete guide to time series forecasting in Python. *Machine Learning Plus*. <https://www.machine-learningplus.com/time-series/arima-model-time-series-forecasting-python/>
- Sengupta, J., Ranjanendra, N. N., & Goswami, B. (2017). Commodity price fluctuations and unemployment in a dependent economy. *Contemporary Economics*, 11(3), 315-326. [10.5709/ce.1897-9254.245](https://doi.org/10.5709/ce.1897-9254.245)
- Seong, B. (2020). Smoothing and forecasting mixed-frequency time series with vector exponential smoothing methods. *Economic Modelling*, 91, 463-468. <https://doi.org/10.1016/j.econmod.2020.06.020>
- Sidqi, F. & Sumitra, I. D. (2019). Forecasting product selling using single exponential smoothing and double exponential smoothing methods. *IOP Conference Series: Material Science Engineering*, 662(3). <https://doi.org/10.1088/1757-899X/662/3/032031>
- Sinković, D., Zemla, S. & Zemla, N. (2021). Monitoring of economic indicators in the context of financial and economic crises. *Contemporary Economics*, 16(1), 61-87. <https://doi.org/10.5709/ce.1897-9254.469>
- Škare, M. & Stjepanović, S. (2016). Measuring business cycles: A review. *Contemporary Economics*, 10(1), 83-94. <https://doi.org/10.5709/ce.1897-9254.200>
- Smigel, L. (2021). How to interpret ARIMA results, *The Analysing Alpha Blog*. <https://analyzingalpha.com/interpret-arima-results>
- Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36(1), 75-85. <https://doi.org/10.1016/j.ijforecast.2019.03.017>
- Stiglitz, J. E. (2012). *Globalization and its discontents* (2nd Ed.). Penguin
- Suraiya, S. & Hasan, M. B. (2018). Identifying an appropriate forecasting technique for predicting future demand: a case study on a private university in Bangladesh. *Dhaka University Journal of Science*, 66(1), 15-19. <https://doi.org/10.3329/dujs.v66i1.54539>
- Taques, F. H., López, M. G., Basso, L. F., & Areal, N. (2021). Indicators used to measure service innovation and manufacturing innovation. *Journal of Innovation & Knowledge*, 6(1), 11-26. <https://doi.org/10.1016/j.jik.2019.12.001>
- Temür, A. S., Akgün, M., & Temür, G. (2019). Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models. *Journal of Business Economics and Management*, 20(5), 920-938. <https://doi.org/10.3846/jbem.2019.10190>

- Trapp, C. T., Kanbach, D. K., & Kraus, S. (2022). Sector coupling and business models towards sustainability: The case of the hydrogen vehicle industry. *Sustainable Technology and Entrepreneurship*, 1(2), 100014. <https://doi.org/10.1016/j.stae.2022.100014>
- Vali, S. (2014). *Principles of mathematical economics*. Atlantis Press. <https://doi.org/10.2991/978-94-6239-036-2>
- Viccione, G., Guarnaccia, C., Mancini, S., & Quartieri, J. (2020). On the use of ARIMA models for short-term water tank levels forecasting. *Water Supply*, 20(3), 787-799. <https://doi.org/10.2166/ws.2019.190>
- Vo, T. T. B. C., Le, P. H., Nguyen, N. T., Nguyen, T. L. T., & Do, N. H. (2021). Demand forecasting and inventory prediction for apparel product using the ARIMA and fuzzy EPQ model. *Journal of Engineering Science and Technology Review*, 14(2), 80-89. <https://doi.org/10.25103/jestr.142.11>
- Wilczyński, R. (2011). Global financial governance: A perspective from the International Monetary Fund, *Contemporary Economics*, 5(1), 4-16. <https://doi.org/10.5709/ce.1897-9254.1>
- Zeng, Q., van Blokland, W. B., Santema, S., & Lodewijks, G. (2019). Benchmarking company performance from economic and environmental perspectives. *Benchmarking: An International Journal*, 27, 1127-1158. <https://doi.org/10.1108/BIJ-05-2019-0223>