Pavelková, Drahomíra; Homolka, Lubor; Vychytilová, Jana et al.

Article

Passenger car sales projections : measuring the accuracy of a sales forecasting model

Ekonomický časopis

Provided in Cooperation with:
Slovak Academy of Sciences, Bratislava

Reference: Pavelková, Drahomíra/Homolka, Lubor et. al. (2018). Passenger car sales projections : measuring the accuracy of a sales forecasting model. In: Ekonomický časopis 66 (3), S. 227 - 249.

This Version is available at:
http://hdl.handle.net/11159/3925

Passenger Car Sales Projections: Measuring the Accuracy of a Sales Forecasting Model[1]

Drahomíra PAVELKOVÁ – Lubor HOMOLKA* – Jana VYCHYTILOVÁ* – Vu Minh NGO* – Le Tuan BACH* – Bruce DEHNING***

Abstract

This paper considers the importance of the automobile industry in the global economic environment and sheds additional insight on the forecasting of passenger car sales. The study uses data from the automotive sectors in 38 countries, which account for more than 80% of passenger cars in use worldwide for testing the accuracy of a general framework that uses income and other country-specific factors to forecast passenger cars sales for short- and mid-term periods. The results indicate that this framework can be applied to a wide range markets, but its performance is primarily influenced by income levels in these markets. Tested and discussed are not only income as the main predictor of sales, but also the effects of other factors such as vehicle ownership level on passenger car sales projections. Income is shown to play both a determining role and a moderating role that affects other variables' impact on passenger car sales.

Keywords: *forecasting model, automotive, passenger car sales, income, accuracy of model, vehicle ownership*

JEL Classification: C52, C53, E23, O18, O50, R41

* Drahomíra PAVELKOVÁ – Lubor HOMOLKA – Jana VYCHYTILOVÁ – Vu Minh NGO – Le Tuan BACH, Tomas Bata University in Zlín, Faculty of Management and Economics, Department of Finance and Accounting, nám. TGM 5555, 760 01 Zlín, Czech Republic; e-mail: pavelkova@utb.cz; homolka@utb.cz; vychytilova@utb.cz; ngominhvu@gmail.com; bachstep1008 @gmail.com

** Bruce DEHNING, Chapman University, Argyros School of Business and Economics, One University Drive, Orange, CA 92866, USA; e-mail: bdehning@chapman.edu

[1] The authors are thankful for the financial support received from the Czech Science Foundation, Grant No. 16-25536S: *Methodology of Developing a Predictive Model of Sector and Company Performance in the Macroeconomic Context.*

Introduction

Transportation is a fundamental prerequisite for a society's development and improvement of people's lives. The automotive sector is considered crucial for global economic development and prosperity – contributing roughly 3% of all GDP output globally. Globally the sector manufactured 91.5 million motor vehicles in 2015 (73.5 million passenger cars). The top three countries in the world, manufacturing more than half of all passenger cars and commercial vehicles, are China (26.99% of the global total), U.S. (13.30%) and Japan (10.22%). Other major motor vehicle producing countries are Germany, South Korea, India, Mexico, Spain, Brazil, Canada, France, Thailand, and the United Kingdom. Cars represent the world's number two export product, surpassed only by crude oil. In 2015, global car exports were valued at USD 672.9 billion. The automotive industry represents the largest private investor in Research and Development (R&D). Globally, almost 6,000 patents are granted to the automotive sector per year. The European Union is by far the world's largest investor in automotive R&D, with EUR 44.7 billion invested per year. In 2013, the automotive sector created 12 million jobs in Europe, more than 8 million in the U.S., and more than 5 million in Japan. The proportion of supplier value added to automobile production increased from 56% in 1985 to 82% in 2015. According to Harrington (2015), this increased dependence in only thirty years changes the hierarchy of players away from the traditional power base where the big original equipment manufacturers held almost all the market power. The automotive sector is important for upstream industries such as steel, chemicals, and textiles, as well as for downstream industries such as Information and Communication Technology (ICT) mainly contributing in adding sophisticated functionalities making a car intelligent; repair and maintenance services for vehicles; or support mobility services. The industry is crucial also by links to the other sectors having important multiplier effects in the economy.[2]

In this paper, we empirically project passenger car sales in and beyond the 2007 – 2008 financial crisis for different countries, which can be important for different reasons. First, forecasting automobile car sales can help automobile companies to better understand their business and use the results for strategic planning. To understand the automotive industry, one needs to understand its historical performance in relation to multiple economic factors that affect the industry. Second, forecasting sales, which requires demand forecasts, is critical for maximizing profits. Forecasts enable an organization to optimize inventory

[2] Data was sourced from various databases, particularly GDP global output from A. T. Kearney; vehicle production, R&D, and patents from ACEA; car exports from WTEx; jobs in the automotive industry from ACEA, JAMA, and the Alliance of Automobile Manufacturers; links between the automotive industry and other sectors from the European Commission.

levels, make appropriate purchasing decisions, and maintain efficient daily operations. Third, forecasting car sales can help companies prepare their operations for fluctuations in the industry in the future (Sanders, 1997; Shahabuddin, 2009). Fourth, growth in vehicle ownership implies growth in the demand for oil (Dargay, Gately and Sommer, 2007). Thus, exploring the topic can also potentially contribute to the field of forecasting long-run gasoline demand for passenger cars (Storchmann, 2005). Finally, the implications of the forecasts generated for each country in the sample may be helpful in solving future transportation and energy-related issues (Medlock and Soligo, 2002).

The goal of this article is to verify correctness usefulness of the *Haugh's model*, developed by Haugh, Mourougane and Chatal (2010). This model is designed to forecast passenger cars sales in a short and mid-term period. Income level expressed as GDP per capita in Purchasing Power Parity (PPP) is used as the main explanatory variable. Some other variables need to be supplied. Therefore, we follow the approach in Dargay, Gately and Sommer (2007) who propose a model to estimate country *vehicle ownership level (stock)* as a function of income and level of urbanization and population density. Their analysed period spanning years 1960 – 2002 accross sample of 45 countries which represent 75% of the world's population. We build on both studies and extend the analysis to more countries and to pre/post financial crisis periods. We also estimate model's parameters on the most recent data, discuss forecasting performance and provide an explanation of the country-level bias.

The four main objectives of this research are to: (i) update variables used in Haugh's model in original and extended set of countries. This includes demographic and automotive data used by Dargay, Gately and Sommer (2007) to re-estimated parameters of the original model; (ii) make a retrospective analysis on three periods (defined later in the text) to assess predictive validity of the model; (iii) evaluate efficacy of the Haugh's model and model with optimised parameters by employing standard forecast error metrics; and to (iv) identify cause(s) of forecast errors.

The results of our research indicate that Haugh's model can be applied to a wide range of car sales markets, but its performance is highly influenced by income levels in those markets and by the current level of market saturation. The relationship between income and passenger car sales might not be fully captured in Haugh's model when the country-level bias is not considered. Additional adjustments can be made to eliminate country-level bias to improve the overall performance of Haugh's model.

The remainder of the paper is organized as follows; after a comprehensive theoretical review in Section 1, the framework for measuring and forecasting passenger car sales and data collection are described in Section 2. In Section 3, results of analysis and hypothesis testing are introduced, followed by a discussion of the findings and conclusions in Section 4. Further research issues are also suggested in the last section.

1. Theoretical Background

Forecasting sales is an essential part of most of business activities. A sales forecast is a prediction based on past sales performance and an analysis of expected market conditions. Many management and control decisions are influenced by the current market situation and on how it is expected to change in the near future. Many decision-making tasks are based on predicted values. As a rule, predictions can only be as good as the quality of input variables (Gahirwal and Vijayalakshmi, 2013; Sanders, 1997). Therefore, prediction errors can arise from both model misspecification and from unreliable underlying data.

In general, there are two main approaches to forecasting: (1) quantitative methods (objective approach), and (2) qualitative methods (subjective approach). While qualitative forecasting techniques employ the judgment of experts in a specified field to generate forecasts, quantitative forecasting methods are based on an analysis of historical data to forecast future data points. Quantitative methods include statistical techniques such as generalised linear regression models, exponential smoothing and auto-regressive modelling strategies, state space models; and machine learning and artificial intelligence techniques such as neural networks or genetic algorithms (Box et al., 2015). The research in this paper focuses on quantitative methods which stand on the sound economic ground.

Researchers and practitioners have developed various quantitative forecasting methods that differ in level of complexity, forecasting accuracy, or in the purpose of modelling. Some of them are embedded into decision-making support system. Considerably large attention was to sales forecasting, see Dalrymple (1987), Fildes and Makridakis (1995), Gahirwal and Vijayalakshmi (2013), Merigo, Palacios-Marques and Ribeiro-Navarrete (2015), Sanders (1997) and more. Dalrymple (1987) conducts a survey to discover how businesses prepare sales forecasts, what methods they prefer, and focuses on the accuracy of their predictions. Thus, Dalrymple (1987) sheds some light on the forecasting methods widely used and the accuracy of their predictions, by surveying 134 executives in the United States. Sanders (1997) extends the study by collecting information about the forecasting techniques, software, and common causes of forecast errors of 350 manufacturing firms in the United States. He also warns that qualitative forecasting methods have been shown in studies to be less accurate than quantitative methods, leading to many biases and therefore contributing to poor quality forecasts. Sanders (1997) points out that many past surveys on forecasting looked at forecasting practices across multiple industries. The author suggests, however, that it is better to focus on firms in a particular sector, for example in his case on manufacturing firms, in an effort to uncover problems specific to this type of industry. Sanders (1997) highlights that between industries there exist differences and requirements on forecasting practices, and thus it is important to understand the characteristics and practices in a specific industry segment.

Fildes and Makridakis (1995) conclude that time series statisticians and empirical researches, if successfully working together, should advance the field to better serve those engaged in decision or policy making. This is accomplished through more accurate predictions and making the forecasting discipline more useful and relevant for real-life applications. Gahirwal and Vijayalakshmi (2013) state that one challenge is to decrease the error of the forecast as much as possible and a second is to find/develop a relatively inexpensive and easy to maintain forecasting system that guarantees the desired accuracy. Merigo, Palacios-Marques and Ribeiro-Navarrete (2015) point out that an important issue in addressing sales forecasts is the calculation of average sales. They add that usually the arithmetic mean or the weighted average are used for such calculations. Furthermore, they propose that the new methods for estimating the average sales using the ordered weighted average, and the unified aggregation operator, can better address uncertain and complex environments.

The literature on forecasting motor vehicle/passenger car sales or ownership varies by the forecasting methods used and in the number of panel countries. Considerable expertise in the economic understanding of forecasting car ownership and use has been developed, particularly with the spur of the petrol price increases of 1973 – 1975 and 1979 – 1981 (Mogridge, 1989). Mogridge (1967) describes a method of short-term forecasting car ownership in the United Kingdom based on the relationship between the expenditure on car purchases and the level of household disposable income.[3] However, this model was only based on the stock-income relation without any mention of household car ownership levels at a given income level or of the park-income relation. Mogridge (1967) defines "stock" as value and uses "park" for a number of cars. A different approach is used in Mogridge (1989), a stock-park relation at a given income level for use when attempting to forecast ownership and use of cars. He adds that UK government long--run forecasts of car ownership have reverted to essentially time-trend forecasting,

[3] Statistics on household's disposable income had become available through the initial Household Expenditure Enquiry of 1953 – 1954 and the subsequent annual Family Expenditure Surveys (FES) since 1959. In 1971, statistics were separated between car-owning households and non-carholding households (Mogridge, 1989).

but with saturation levels of car ownership that are likely far too low. Similarly, Dargay, Gately and Sommer (2007) point out the analyses of International Energy Agency (IEA) and Organization of the Petroleum Exporting Countries (OPEC) contain assumptions about vehicle saturation rates which are much lower than actual vehicle ownership already experienced in most higher-income countries.

In the major part of the literature, demand for motor vehicles is a function of income (for example Button, Ngoe, and Hine, 1993; Dargay and Gately, 1999; Dargay, Gately and Sommer, 2007; Haugh, Mourougane and Chatal, 2010; Mogridge, 1967 and more). Button, Ngoe, and Hine (1993) forecast car park for 10 low-income countries in 2000 and 2025 using a quasi-logistic car ownership model, and define "car park" as the total number of vehicles in the country. The main explanatory variable in the model influencing per capita vehicle ownership at the national level is income. They point out that increase in the overall car park moves ahead of the rate of increase in per capita ownership as populations expand. Medlock and Soligo (2002) examine the effect of economic development on the demand for private motor vehicles for a panel of 28 countries. They develop a model of the relationship between economic development and per capita private car ownership. They find that saturation levels vary across countries and that users' costs are a significant factor in the evolution of vehicle stocks.

Dargay and Gately (1999) were the first to include countries for dynamically specified[4] model estimation over the period 1960 – 1992 covering the full range of income levels. Their study projects growth in the car and total vehicle stock until 2015, based on an econometrically estimated model explaining the growth of the car/population ratio (referred as car ownership) as a function of per-capita income and estimated demand in a sample of 20 OECD countries and six developing countries. Dargay, Gately and Sommer (2007) make a significant contribution to the topic by explaining historical patterns in vehicle ownership rates as an S-shaped Gompertz function of per-capita income and covering 45 countries over the period over 1960 – 2002. Their model exploits similarity of response in vehicle ownership rates to per-capita income across countries over time while allowing for cross-country variation in the speed of vehicle ownership growth and in ownership saturation levels. They point out the relationship between vehicle ownership and per-capita income is highly non-linear. At very high levels of income, vehicle ownership growth decelerates and slowly approaches the saturation level, this was the case in most OECD countries in 2007, the year of publishing the paper. Haugh, Mourougane and Chatal (2010) build on Dargay,

[4] Incorporating short- and long-run income elasticity of car and vehicle ownership dependent upon per-capita income. The elasticity ranges from about 2.0, for low- and middle-income levels down to zero for the highest income level. At low and middle-income levels, ownership grows twice as fast as income and ownership saturation exists at the highest income level (Dargay and Gately, 1999).

Gately and Sommer (2007) and derive short and mid-term projections of car sales and consider the role of the automobile industry in the current cycle. Among other things, they highlight the importance of the automotive industry and how its cycle intertwines with the business cycle. Previously Ramey and Vine (2005) had shown that the automobile industry experiences decline in production around the same time as the rest of the economy. They show that declines in the automobile industry were even more dramatic than overall economic declines.

Haugh, Mourougane and Chatal (2010) show that trends in mid-term projections of car sales in mature markets such as Europe and North America are likely to remain stagnant. On the other hand, they foresee rapid increases in trends of sales in China and in the five largest Western European countries. Their forecast for automakers in the NAFTA area points to decline in domestic market share and to increased reliance on exports to avoid excess capacity. The authors conclude that in the medium-term car manufacturers will face different demand conditions around the world. Comparing trend sales with production capacity can provide perspective on the forces producers in various countries may be facing. Whether manufacturers have excess capacity in each country or area depends on their ability to compete for market share in their home market and in export markets.

The developing literature on forecasting car sales in the automotive industry varies mainly in types of forecasting approaches applied (for more see for example Wu, 2009, who uses a forecasting model based on a wavelet v-support vector machine; Landwehr, Labroo and Herrmann, 2011, who incorporate design fluency as a predictor in an automotive car sales forecast; Daly and Gallachóir, 2011, who model future private car energy demand, taking into account the lifetime survival profile of different car types, the trends in vehicle activity over the fleet, and the fuel price and income elasticity of new car sales and total fleet activity; and more). Other authors attempt to forecast motor vehicle sales or ownership requiring a minimal set of assumptions about demographic, economic and demand trends (Homareau, 2015; Greenspan and Cohen, 1999; Pierdzioch, Rülke and Stadtmann, 2011).

Barbera, Clickb and Darraough (1999) analyse the impact of changes in exchange rates and oil-prices on the market-shares of Japanese and American automakers in the U.S. market. They point out that an appreciation of the Yen increases the quantities sold by American automakers and decreases the quantities sold by Japanese automakers. Also, an increase in oil-price reduces the number of automobiles sold by American automakers but it has only a small effect on Japanese automakers. Shahabuddin (2009) used regression analysis to obtain highly correlated automobile sales, economic, and demographic variables. The results indicate a strong relationship between multiple economic variables and foreign car sales. Sharma and Sinha (2012) use a fuzzy neural back propagation algorithm (BPN algorithm) to implement a sales forecast model of the automobile industry in in India for Maruti Suzuki Ltd. The inflation rate, petrol price, and previous month sales are found to be the most significant factors influencing the car sales forecast for this company.

Other forecasting techniques belong to studies by Dwivedi, Niranjan, and Sahu (2013), or Sa-Ngasoongsong et al. (2012). Dwivedi, Niranjan, and Sahu (2013) use two forecasting methods, moving average and exponential smoothing, to forecast past sales in the automobile industry and then use the forecasted values as an input for an adaptive neuro fuzzy inference system (ANFIS) to obtain the final sales forecast. They conclude, after comparing the results with two other forecasting models, namely artificial neural network (ANN) and linear regression, the empirical results favour the ANFIS model over the other two. Sa-Ngasoongsong et al. (2012) indicate that automobile sales at the segment level have a long-run equilibrium relationship (cointegration) with the identified economic indicators. They estimate a vector error correction model (VECM) of multi-segment automobile sales based on impulse response functions to quantify the long-term impact of these economic indicators on sales. Comparisons of prediction accuracy demonstrate that the VECM model outperforms other classical and advanced time-series techniques. The empirical results suggest that the VECM can significantly improve the accuracy of predicting automotive sales for a 12-month ahead prediction in terms of Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE), compared to standard time series techniques.

2. Methodology

In this paper, we build on the Haugh, Mourougane and Chatal (2010) forecasting approach, which incorporates Dargay, Gately and Sommer (2007) estimations. Using the framework of Haugh, Mourougane and Chatal (2010), the relationship between income and passenger car sales is presented. Inspired by the aim of discovering the predictive power of income (using GDP in PPP per capita as a proxy) on passenger car sales, the evaluation of Haugh's model's ability to predict passenger car sales can give us some new insights. This investigation provides us with the background to understand the real impact of income on passenger car sales and then to improve Haugh's model's predictive power.

2.1. Framework for Measuring Passenger Car Sales

According to Haugh, Mourougane and Chatal (2010), the number of passenger cars sold ($sales_{it}$) is equal to the sum of the number of scrapped passenger cars ($scrappage_{it}$) and the change in the passenger car stock ($\Delta stock_{it}$):

$$sales_{it} = \Delta stock_{it} + scrappage_{it} \tag{1}$$

The estimate of scrapped passenger cars is the product of the historical average scrap rate (asr_i) and the prior year's passenger car stock ($stock_{it-1}$): $scrappage_{it} = asr_i * stock_{it-1}$, where the historical average scrap rate (asr_i) is computed as:

$$asr_i = \sum_{t=1}^{T} \frac{sale_{it} - \Delta stock_{it}}{stock_{it}} \tag{2}$$

The predicted passenger car stock ($stock_{it}$) is determined by the per capita passenger car stock (pc_{it}) multiplied by the total population (pop_{it}):

$$stock_{it} = pc_{it} * pop_{it} \tag{3}$$

The per capita passenger car stock (pc_{it}) depends on the historical average of the passenger car stock to total vehicle stock ratio (pcr_i) and per capita vehicle stock (v_{it}):

$$pc_{it} = pcr_i * v_{it} \tag{4}$$

The determination of per capita vehicle stock is based on the previous year's per capita vehicle stock (v_{it-1}) and the adjustment of short-term trend in vehicle ownership to the long-term equilibrium level, $\theta(vlr_{it} - v_{it-1})$:

$$v_{it} = v_{it-1} + \theta(vlr_{it} - v_{it-1}) \tag{5}$$

where

θ – the speed of adjustment and

vlr_{it} – the long-term equilibrium per capita vehicle stock obtained by a Gompertz function:

$$vlr_{it} = \gamma_i e^{\alpha e^{\beta_i GDP_t}} \tag{6}$$

where

γ_i – the saturation level of per capita vehicle stock,

α and β_i – define the curvature of the function.

Gompertz function is selected to describe the long-term relationship between the vehicle ownership and per capita income (GDP_t) in an S-shaped curve due to its dominant simplicity and flexibility. Dargay, Gately and Sommer (2007) include a Gompertz function in the vehicle stock estimating model and determined the values of θ, γ_i, α and β_i. Haugh, Mourougane and Chatal (2010), in turn, use these estimated parameters as the previously determined values in the passenger car sales model.

We evaluate Haugh's model on three periods. In the period 1995 – 2002 (Period I) all parameters estimated by Dargay, Gately and Sommer (2007) are used. This can be viewed as an in-sample analysis as these parameters were identified in the period 1960 – 2002. Second period (Period II) ranges from 2003 – 2009. Parameters remained unchanged but the analysis is now out of the sample. Last period (Period III) end in 2015. In the Period III selected parameters of Gompertz curve are re-estimated by minimising square errors of the period 1995 – 2009. This model is hereafter called *Optimised model*.

Two models will be considered in this paper. The *Original* model with Dargay's estimates and Optimised model. The *Optimised* model was built by minimising MSE on the Period I and II. Parameters were identified by using non-linear Generalized Reduced Gradient (GRG) method. We have optimised several parameters simultaneously: α, β, ϕ, while keeping original values of saturation levels. It was necessary to update starting value of short-term trend of vehicles per capita in 1995. This value was computed from average growth in the phases I and II. Haugh, Mourougane and Chatal (2010) optimised β values, too. Their objective function was set as a target range $\pm 2.5\%$ of total real sales. They optimised β values while keeping other parameters as in Dargay's paper so the sum of estimated sales falls into the target range. We tried to replicate this approach as well. Unfortunately for many countries we failed to find a feasible solution, or some derived parameters were unjustifiably low/high (such as a proportion of passenger cars to total cars low as 5%).

MAPE is used for evaluating the forecast errors because it ignores the scale, which affects other error measures like Mean Absolute Difference (MAD) or MSE. Thus, MAPE is suitable to study the forecast errors of different countries regarding their differences in volumes of passenger car sales.

2.2. Data

First, definitions of economic variables and their sources are defined (see Appendix A). In total, data from 38 countries are collected including most of the world's biggest economies from different geographical regions around the world. These economies are leaders in consumption and production of passenger cars. We analyse 38 countries which account for more than 80% of passenger cars in use in 2014 (OICA, 2017).[5] Using the framework of Haugh, Mourougane and Chatal (2010), income (GDP in PPP per capita as a proxy) is used as the main macro indicator determining passenger car sales. GDP data provided by World comparable outcomes. In case of few countries we have not received identical values

[5] In an attempt to test the feasible range of Haugh's model, we extend the sample of countries from the original 17 in Haugh, Mourougane and Chatal (2010) to 38 countries.

of GDP Bank were transformed to 1995 prices to match Dargay's underlying data to obtain and car sales to Dargay's, which we attribute to reviews occurred after initial publication (such as deflator values). Few manual adjustments reflecting unique conditions were made. For example, we have smoothed time series of scrappage rates in Germany to eliminate a sharp decline of the total amount of car sales in 2007 (first shock wave of economic crisis) and the effect of scrappage scheme in 2009. These extraordinary events do not reflect long-term conditions in ordinary years. Extreme values remained in the model for particular years and inflated errors, though.

We evaluate Haugh's model performance in three periods. 13 countries had a complete history of the period I and II. There were three countries with seven years, 18 countries with five years and one country with only four years history. All countries were recorded over six years in Period III.

2.3. Tested Hypothesis and Processing Methods

This subsection presents overview of hypotheses designed to assess a quality of the Haugh and our optimised model. Hypotheses were stated in a form which expects validity of the economic models. That means we postulate no difference between groups or unbiased prediction (absence of bias). Empirical analysis aims to uncover whether these expectations are met. Our restricted data sample does not allow us to employ the null hypothesis statistical hypothesis testing in most of the cases. However, if such methods would have been used, an alternative hypothesis would be in line with what we have done by analysis of errors (e.g., finding an evidence for an existence of bias.)

First two hypotheses concern an ability of Haugh's and our optimised model to capture growth rates across countries.

H1A: *Haugh's model's projected growth rates of passenger car sale does not differ to the real growth rates in the period 2010 – 2015.*

H1B: *Projected growth rates of Haugh's model with optimised parameters do not differ to the real growth passenger car sales rates in the period 2010 – 2015.*

Our next concern is to assess unbiasedness of the model. We consider model as unbiased if the expected value of forecast error equals to 0.

H1C: *Mean value of forecast errors of Haugh's models (original and updated) of passenger car sales in the period 2010 to 2015 is zero.*

After the general features of the model is probed we move to an assessment of the improvement of the optimised model to original model. We also analyse models' stability in time and with respect to varying national level of income. Therefore, we propose following hypothesis:

H2A: *Forecasting errors in the optimised model have lower value of MAPE.*

H2B: *Forecast errors of Haugh's of one type are not significantly different between the 2005 – 2009 and 2010 – 2015 period.*

H2C: *Forecast errors of Haugh's models do not significantly differ between countries with low and high values of income per capita.*

Methods and techniques used in the process of hypotheses verification are described in corresponding sections as they are part of the reasoning about the studied issues.

3. Results of Analysis

Forecasting errors are derived by comparing point projections to the real passenger car sales of 38 countries. This step estimates effects of shocks resulting from the change of income on the passenger car sales which cannot be explained by Haugh's and Optimised model. Forecast errors are also used in the parameters' optimisation.

Haugh's model aims at estimating both values in levels as well as general trends across countries in the short and medium term. Concerning to the ability of Haugh's model to produce the unbiased estimated of the passenger car sale trend in medium term across different countries, there are two criteria which Haugh's model should be investigated. First, Haugh's model should be able to replicate passenger car sales growth. Sales growth rates should be insignificantly different from the real growth rates of passenger car sales. If the real and predicted growth rates are reasonably similar, the model can predict general tendency of the time-series. Second, forecast errors should be unbiased. In the best case, Haugh's model forecast errors on test sample (Period III) should be as small as possible; ideally just include irreducible error which is immanent to the studied phenomena stripped of the model's errors. Some of Haugh's model characteristics were identified on general sample (α or max. saturation level), some parameters are estimated to capture country-specific factors (β). To evaluate Haugh's model predictive ability, forecast errors of total passenger car sale from 2010 to 2015 for each country are calculated. If the prediction is unbiased, mean value of errors should be zero or close to zero. As the test-period is short, only 5 years, we do not formally test indifference from 0 by statistical test (such as t-test or Wilcox test). Instead, we provide graphical evaluation accompanied with descriptive statistics. To demonstrate Haugh's model ability to capture a trend, following hypotheses are stated:

H1A: *Haugh's model's projected growth rates of passenger car sale does not differ to the real growth rates in the period 2010 – 2015.*

H1B: *Projected growth rates of Haugh's model with updated parameters do not differ to the real growth passenger car sales rates in the Period 2010 – 2015.*

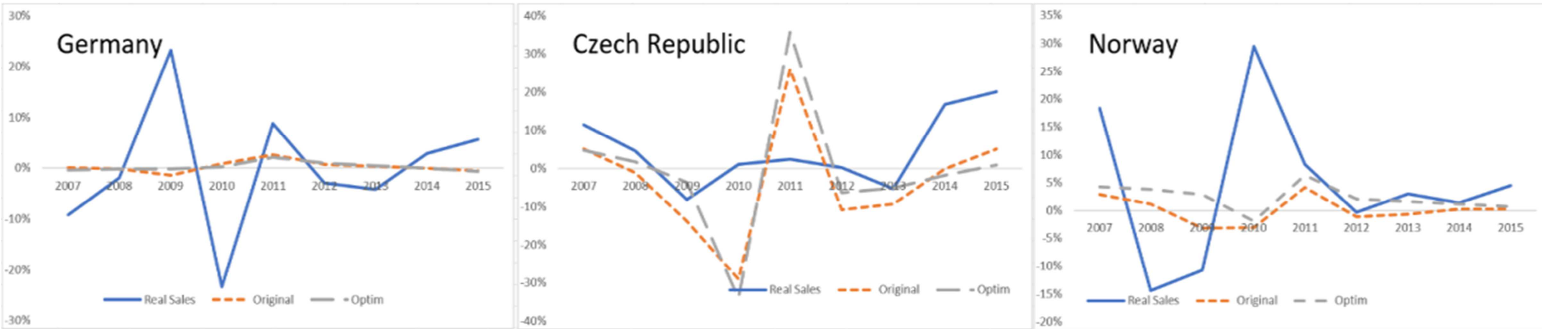
To address hypothesis H1A and H1B, we need to perform a country-level analysis. Whether the form of growth is identical or not is decided by visual inspection. On some countries predicted growth rates from (i) both models were aligned with real growth rates. To this category belongs Argentina, Brazil, Chile, China, Denmark, Finland, Indonesia, Ireland, Netherlands, Norway, Turkey, U.S., South Africa and Czech Republic (except one year), (ii) countries when only one model mimics growth rates (Egypt, Greece, Hungary, Spain, Sweden), (iii) countries when model smooths the time series (Canada, France, Germany, Israel, Japan, Mexico, Switzerland, United Kingdom) and finally (iv) countries where the trend is not captured by any model. To this class usually belong countries where predicted values are lagged to the real one (Australia, Korea, Morocco) or when the trend is not captured at all (Austria, Belgium, Ecuador, India, Italy, Pakistan, Poland, Thailand).

Models of Germany and France smoothed time series over shocks, such as scrappage programs. In Australia, where the number of cars is further away from saturation level, model was more sensitive to changes in GDP. This reaction to Income change is partially reflected in the current year, but mostly is propagated to the following period.

To demonstrate our results and classification, we've selected three countries and its estimates. Although analysis of growth (by MAPE) is conducted on the period 2010 – 2015, predeceasing periods are displayed in the figures to capture dynamics of the development in Figure 1. Germany was also selected as a main European car producer. The sharp decline in auto sales in Germany in 2007 can be explained in several ways. Initially, OECD (2009) states that the shift of the production towards non-OECD regions could result in a decrease in the share of the global production. For example, while vehicle sales have about doubled in China, India, Thailand, and Indonesia since 2007, vehicle markets in the EU remained constant or decreased within the same time period as published by International Council on Clean Transportation (ICCT, 2013). Therefore, the economic crisis may serve to reinforce and accelerate this trend (Haugh, Mourougane and Chatal, 2010). Furthermore, in response to the economic crisis, all the German auto firms had a massive cut in production and even consider the closure of entire factories (Weller, 2008). This reaction seems to be stronger than other countries, which leads to a deeper decline in auto sales than other countries. Strong growth in 2009 was caused by the governmental scrappage program. Drop in 2010 is caused by ending scrappage scheme. In the recent years, car sales are gaining buying momentum, which is not reflected by both models. Both models are on the mean-reverting trajectory. This is caused by market saturation, which, according to the model, is not able to absorb more car despite GDP growth.

F i g u r e 1

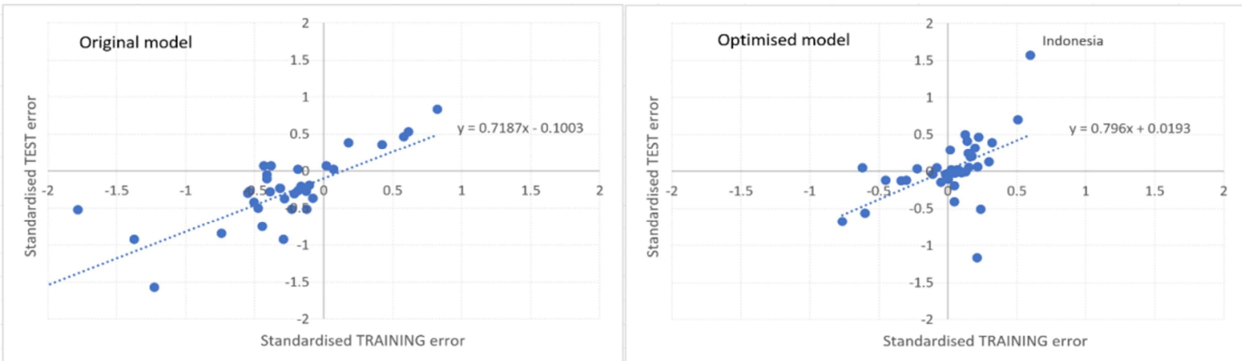
Real Sales (solid line) and Estimated Growth Characteristics for Original Model (dotted) and Optimised Model (dashed)



Source: Own processing.

F i g u r e 2

Comparison of Mean Values of Two Models



Source: Own processing.

Economics of the Czech Republic strongly depends on the performance of Germany's economy. Compared to Germany we see neither a huge dip in 2007 nor effect of scrappage program in Germany. Both models are correctly identifying acceleration of car sales growth.

The last selected case is Norway. It's because it has the lowest MAPE on the Period III on the Optimised model, but it's only 21st best performing on the Original model. This discrepancy is caused by the large value of real growth in 2010 which was also recorded in other Northern European countries. The optimised model has adjusted better to this growth-pattern, while the Original model expected growth rate around 0%.

To assess a bias of predictions, we postulate following hypothesis:

H1C: *Mean value of forecast errors of Haugh's models (original and updated) of passenger car sales in the period 2010 to 2015 is zero.*

We do not directly test null hypothesis as we were constrained by a small number of observations. Instead, we compute the standardised mean value of an error that is an average error to average sales in the period.

Closer analysis of standardised errors reveals that magnitude of errors on the training set (Periods I and II) and test set (Period III) reveals that mean values in both samples tend to have the same sign, as depicted in Figure 2. This indicates that most of the models are biased in the same direction.

The symmetry of the errors can be explained by country-level bias in level value, which is not directly modelled in Haugh's model. This might not be a problem in a practical application where manual adjustments can be done by adding a constant to the model. As an example can serve U.S. and Greece. In case of U.S. both models systematically overestimate real sales, in case of Greece underestimate.

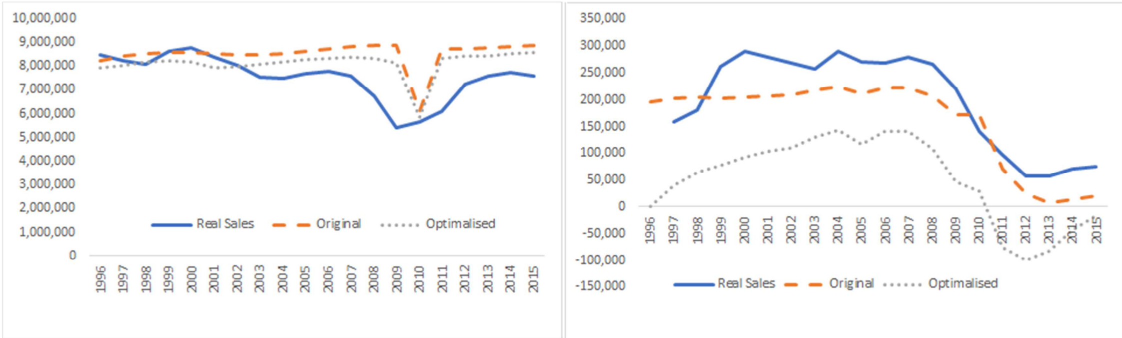
There are 16 countries for which standardised error was smaller than ±10% in the Optimised model. There are only 7 in the original model. As seen from Figure 4 standardised errors are centred around 0 as expected by H1C. Judging from the all-countries analysis, both models provide unbiased predictions. However, as noted above, this doesn't hold true at the country level.

H2A: *Forecast errors in the Optimised model have lower value of MAPE.*

Performance of two models was assessed by comparing MAPE values on the testing set. Figure 5 suggests that Optimised model lowered MAPE values substantially. This was supported by the Mann-Whitney test ($V = 616$, $p\text{-value} < 0.01$) which also identified the difference of mode values. The difference in mean MAPE values is 19.7%. This reduction was not seen in all countries. As Table 1 shows, in some countries (most notably Greece and Norway) MAPE has even increased.

F i g u r e 3

Estimated Values of Passenger Car Sales in U.S. (left) and in Greece, Compared to Real Sales



Source: Own processing.

F i g u r e 4

Histogram of Standardised Mean Values with the same bin width = 0.3. Egypt, Hungary and Pakistan were Removed from the Original Model Due to their High Values Exceeding –1.5

Source: Own processing.

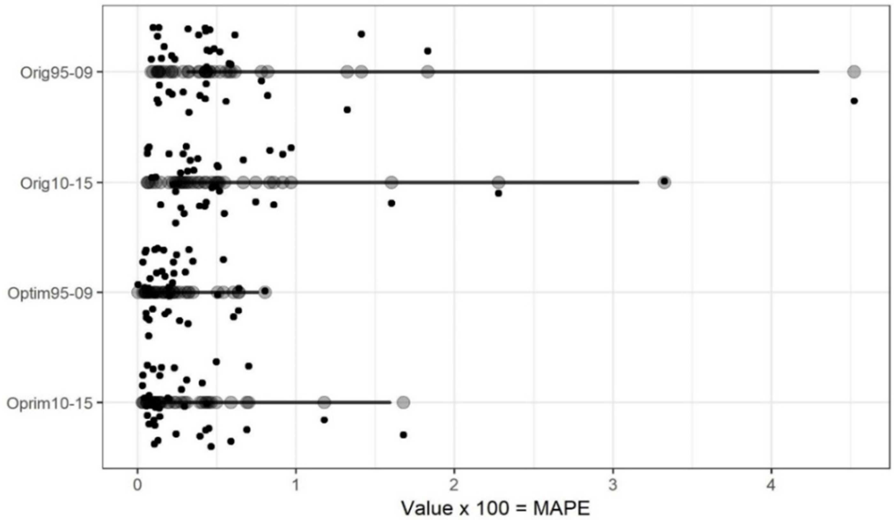
Another important feature of the predictive model is a stability of forecast errors.

H2B: *Forcast errors of Haugh's of one type are not significantly different between the 2005 – 2009 and 2010 – 2015 period.*

Our data exhibit large and unequal variances in groups. Therefore, we employ non-parametric paired Mann-Whitney test to verify whether the changes of mode values of MAPE differ.

F i g u r e 5

Performance Comparison of Two Models on Training and Testing Set



Source: Own processing.

We compare differences of mode MAPE values in the training and testing sample. We perform this test only for model with original parameters. For optimised model first two periods were used as a set of values for objective function in parameters identification so it should not be surprising that the error is substantially smaller. The mode value of MAPE in training set for the Original model was 42,7% while for test sample 36,8%. This small difference is not conclusive as the abovementioned test failed to reject null hypothesis of equal mode values (V = 366, p-value = 0.5286). For the Optimised model, we have 17.3% MAPE value on the training and 17% on the testing set. We can conclude that, on the general level, there is no evidence that forecasting performance differs. We can consider errors to be stable over all three periods.

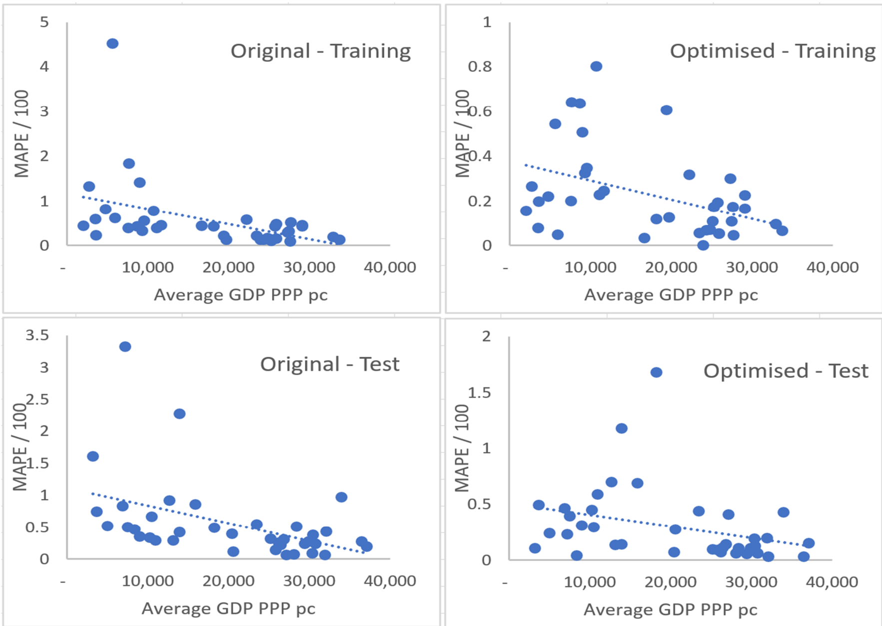
Concerning the applicable range of a predictive model, projected results should be consistent over time and across different contexts to overcome the over-fitting issue. If it is the case, the model is overall reliable and the parameters used in the model are not biased for any specific period of time or particular type of market. Haugh's model is one of the general frameworks that describe the relationships between income and passenger car sales. Therefore, prediction results should be insignificantly different when applying the model in a different period of time and contexts. For testing the applicability of Haugh's model, not just the original sample of countries in Haugh, Mourougane and Chatal (2010) are used but also several other countries are involved into the investigation. The original sample of countries consists of 17 countries and mainly focus on OECD countries plus some big economies such as India, China. The extended sample of countries will consist of other 21 countries that also include countries from other regions ignored in the original samples such as Africa, Middle East, and Oceania. Thus, the extended sample consists of car sale markets that have different characteristics with the ones in original sample. Following hypothesis should be verified:

H2C: *Forecast errors of Haugh's models do not significantly differ between countries with low and high values of Income per capita.*

This hypothesis will be addressed in Figure 6. Linear regression fit in all four windows indicates that with increasing income per capita model results in lower MAPE errors.

Figure 6

Scatter Plot of Average Value of MAPE and GDP per capita in PPP



Source: Own processing.

List of 10 countries with best forecast and 10 worst forecast errors can be found in the following Table 1. The total score is computed as a sum of rank of MAPE errors obtained on test sample from both models. Country with smallest MAPE is assigned number 1, the worst performing number 38.

T a b l e 1

List of Countries with the Best (left part) and Worst Performing Models

Country	Total Score	Rank Original + Optimised	GDP rank	MAPE [%]	Country	Total Score	Rank Original + Optimised	GDP rank	MAPE [%]
Finland	8	3 + 5	12	7.1; 6.0	Thailand	55	30 + 25	29	66.8; 29.5
Australia	13	9 + 4	10	23.9; 5.6	Italy	59	29 + 30	19	54.7; 43.9
Switzerland	14	12 + 2	2	27.2; 3.2	Egypt	60	38 + 22	34	332.5; 23.1
Korea	15	6 + 9	17	14.6; 9.0	Greece	63	25 + 38	22	49.4; 167.9
Austria	16	10 + 6	6	24.1; 6.0	Ireland	64	35 + 29	3	97.0; 42.9
UK	19	11 + 8	16	25.6; 7.2	Indonesia	64	32 + 32	35	83.4; 46.4
Norway	22	1 + 21	5	6.1; 19.8	India	64	31 + 33	37	74.4; 49.7
Japan	22	8 + 14	15	22.3; 11.2	Poland	68	33 + 35	23	85.9; 68.9
Sweden	24	23 + 1	4	43.1; 3.2	Turkey	70	34 + 36	27	91.6; 70.1
Netherlands	24	4 + 20	8	9.0; 18.9	Hungary	74	37 + 37	25	227.8; 117.9

Note: The total score is a sum of ranks of Original and Optimised model. Norway model has had the best predictive performance in the original model, but 23rd best performance on the Optimised. GDP column ranks countries from the highest GDP PPP pc. The last column presents MAPE value on the test sample. There were 31 of 38 countries for which MAPE on the optimised model has decreased compared to the original model.

Source: Own processing.

4. Discussion and Conclusion

According to the results presented in previous sections, some findings can be formulated. Haugh's model can be used to produce consistent and unbiased passenger car sales forecasts for most of the countries if the manual adjustments are made. We note that predictive performance varies as the level of GDP per capita changes. Car sales are more difficult to predict in less developed countries, even model is tailored to a particular country (optimised parameters).

To understand the impact of income on passenger car sales forecasting results, the investigation moves its attention to other factors. Other factors affecting predictions is the current vehicle ownership, saturation level of vehicle ownership in the country, the beta (β) parameters which govern the income elasticity on vehicle ownership and the theta (θ) which control the projection of vehicle ownership. According to previous findings, with the beta (β) unadjusted, the optimal theta (θ) for each group of countries to control the impacts of income level on the projection of passenger car sales is different from each other. At the lower income countries, because of the low level of the current vehicle in use, the difference between the long-term vehicle ownership and the current vehicle in use are very high. Thus, the lower theta (θ) might be better to project the vehicle in use, avoiding the overestimation issue. Therefore, using different theta (θ) between lower income country group and higher income country group is more likely to improve the Haugh's model accuracy.

Haugh's model uses GDP as an explanatory variable. This value at time t has to be supplied to estimate passenger car sales at time t. In the paper, we've assumed that analyst knows such GDP for the current year. However, this value is unknown. We've followed this approach to remove uncertainty of GDP estimate to better capture model's uncertainty. All results are therefore the best possible as real but unknown values were used.

We propose some adjustments to the original studies, which focus on GDP as the main driver of car sales. Firstly, more macro variables or sectors variables should be added to fully capture and control the impacts of income on passenger car sales. Several assumptions about shared characteristics across countries are usually made although countries experience different phases of the economic cycle or level of development. For lower income countries, the increase in income might not lead to the rise in passenger car purchases as fast as in higher income countries. Therefore, using only income for projecting passenger car sales might lead to the bias. Secondly, although Haugh's model shows its potential in projecting the passenger car sales in short and medium terms, the projection accuracy is too much dependent on the selection of parameters. Therefore, further research should consider involving more factors to the model to mitigate the role of parameters and hence, making the model of forecasting more reliable and interpretable. This approach will also contribute much more value to the decision-making process of government agencies and firms in automobiles industry.

For those countries where the current vehicle in use (v_t) is already high, like in Germany, U.S., UK or Northern Europe and Western Europe countries, new sales are highly influenced by the scrappage rate. Therefore, for countries in this group where the scrappage rate is stable and does not significantly deviate from the average value errors will be reasonably small. Therefore, in the future model, the prediction of scrappage rate seems to be a promising step.

This paper also has some limitations. First, although all the data used in this paper are collected from official sources, some data was not identical to those in the original paper. We suspect that revisions occurred after the original paper was published. We also needed to do some small adjustment to eliminate outlier. The quality of the data could contribute to the big prediction errors in some countries. Second, the parameters from Dargay should be re-estimated to reflex the current trend of the automotive industry and the new approach suggested in this paper. Third, some parts of the analysis are based on visual inspection of the time series plots which is affected by subjective stands of authors.

We failed to estimate β parameters using Haugh's approach with range $\pm 2.5\%$ of total sum of real sale values. It resulted in unreasonable low values of either short-term trend vehicle ownership per capita, or proportion of passenger cars to all vehicles, which we needed to estimate as a starting point of our optimisation.

References

BARBERA, R. – CLICKB, R.W – DARRAOUGH, M. (1999): The Impact of Shocks to Exchange Rates and Oil Prices on U.S. Sales of American and Japanese Automakers. Japan and the World Economy, *11*, No. 1, pp. 57 – 93.

BOX, G. E. – JENKINS, G. M. – REINSEL, G. C. – LJUNG, G. M. (2015): Time Series Analysis: Forecasting and Control. 5th ed. New York: Wiley.

BUTTON, K. – NGOE, N. – HINE, J. (1993): Modelling Vehicle Ownership and Use in Low Income Countries. Journal of Transport Economics and Policy, *27*, No. 1, pp. 51 – 67.

DALRYMPLE, D. J. (1987): Sales Forecasting Practices: Results from a United States Survey. International Journal of Forecasting, *3*, No. 3 – 4, pp. 379 – 391.

DALY, H. E. – GALLACHÓIR, B. P. (2011): Modelling Future Private Car Energy Demand in Ireland. Energy Policy, *39*, No. 12, pp. 7815 – 7824.

DARGAY, J. – GATELY, D. – SOMMER, M. (2007): Vehicle Ownership and Income Growth Worldwide: 1960 – 2030. The Energy Journal, *28*, No. 4, pp. 143 – 170.

DARGAY, J. – GATELY, D. (1999): Income's Effect on Car and Vehicle Ownership Worldwide: 1960 – 2015. Journal of Transport Economics and Policy, *33*, No. 2, pp. 101 – 138.

DWIVEDI, A. – NIRANJAN, M. – SAHU, K. (2013): A Business Intelligence Technique for Forecasting the Automobile Sales Using Adaptive Intelligent Systems (ANFIS and ANN). International Journal of Computer Applications, *74*, No. 9, pp. 7 – 13.

FILDES, R. – MAKRIDAKIS, S. (1995): The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting. International Statistical Review, *63*, No. 3, pp. 289 – 308.

GAHIRWAL, M. – VIJAYALAKSHMI, M. (2013): Inter Time Series Sales Forecasting. International Journal of Advanced Studies in Computers, Science and Engineering, *2*, No. 1, pp. 55 – 66.

GREENSPAN, A. – COHEN, D. (1999): Motor Vehicle Stocks, Scrap Page, and Sales. The Review of Economics and Statistics, *81*, No. 3, pp. 369 – 383.

HARRINGTON, L. (2015): The Quiet Revolution: Convergence and the Future Automotive Supply Chain. [DHL Supply Chain White Paper.] [On-line.]

HAUGH, D. – MOUROUGANE, A. – CHATAL, O. (2010). The Automobile Industry in and Beyond the Crisis. [OECD Economics Department Working Papers, No. 745.] Paris: OECD.

HOMAREAU, J. (2015): Forecasting Sales Volumes with Economic Indicators. The Journal of Business Forecasting, *34*, No. 3, pp. 32 – 34.

ICCT (2013): European Vehicle Market Statistics Pocketbook 2013. Retrieved September 2017. Available at: <http://www.theicct.org/sites/default/files/publications/EU_vehiclemarket_pocketbook_2013_.pdf>.

LANDWEHR, J. R. – LABROO, A. A. – HERRMANN, A. (2011): Gut Linking for the Ordinary: Incorporating Design Fluency Improves Automobile Sales Forecast. Marketing Science, *30*, No. 3, pp. 416 – 429.

MEDLOCK III, K. B. – SOLIGO, R. (2002): Automobile Ownership and Economic Development: Forecasting Vehicle Demand to the Year 2015. Journal of Transport Economics and Policy, *36*, No. 2, pp. 163 – 188.

MERIGO, J. M. – PALACIOS-MARQUES, D. – RIBEIRO-NAVARRETE, B. (2015): Aggregation Systems for Sales Forecasting. Journal of Business Research, *68*, No. 11, pp. 2299 – 2304.
MOGRIDGE, M. J. H. (1967): The Prediction of Car Ownership. Journal of Transport Economics and Policy, *1*, No. 1, pp. 52 – 74.
MOGRIDGE, M. J. H. (1989): The Prediction of Car Ownership and Use Revisited: The Beginning of the End? Journal of Transport Economics and Policy, *23*, No. 1, pp. 55 – 74.
OECD (2009): Responding to the Economic Crisis: Fostering Industrial Restructuring and Renewal. Retrieved September 2017. Paris: OECD. Available at: <https://www.oecd.org/sti/ind/43387209.pdf>.
OICA (2017): Production Statistics. (n.d.). Retrieved April 19, 2016, Available at: <http://www.oica.net/category/production-statistics/2014-statistics/>.
PIERDZIOCH, CH. – RÜLKE, J. CH. – STADTMANN, G. (2011): Forecasting U.S. Car Sales and Car Registrations in Japan: Rationality, Accuracy and Herding. Japan and the World Economy, *23*, No. 4, pp. 253 – 258.
RAMEY, V. – VINE, D. (2005): Tracking the Source of the Decline in GDP Volatility: An Analysis of the Automobile Industry. [Finance and Economic Discussion Series, No. 14.] Washington, DC: Federal Reserve Board.
SANDERS, N. R. (1997): The Status of Forecasting in Manufacturing Firms. Production and Inventory Management Journal, *5*, No. 2, pp. 32 – 36.
SA-NGASOONGSONG, A. – BUKKAPATNAM, S. T. S. – KIM, J. – IYER, P. S. – SURESH, R. P. (2012): Multi-step Sales Forecasting in Automotive Industry Based on Structural Relationship Identification. International Journal of Production Economics, *140*, No. 2, pp. 875 – 887.
SHAHABUDDIN, S. (2009): Forecasting Automobile Sales. Management Research News, *32*, No. 7, pp. 670 – 682.
SHARMA, R. – SINHA, A. K. (2012): Sales Forecast of an Automobile Industry. International Journal of Computer Applications, *53*, No. 12, pp. 0975 – 8887.
STORCHMANN, K. (2005): Long-run Gasoline Demand for Passenger Cars: The Role of Income Distribution. Energy Economics, *27*, No. 1, pp. 25 – 58.
WELLER, L. (2008): Crisis Grips German Auto Industry. Retrieved September 2017, from International Committee of the Fourth International (ICFI). Available at: <https://www.wsws.org/en/articles/2008/12/germ-d06.html>.
WU, Q. (2009): The Forecasting Model Based on Wavelet v-support vector Machine. Expert Systems with Applications, *36*, No. 4, pp. 7604 – 7610.

Appendix A

Variables Required and their Sources

No.	Variables	Definition	Data sources from	Note
1.	*pop(t)*	Total population at time (*t*) in the country	United Nation. The data is collected from 2005 – 2015	World Population Prospects: The 2015 Revision
2.	*population density (λ),*	Calculated by diving total population by land area (square kilometre)	World bank's World Development indicators database from 2003 – 2015	Used for re-estimate beta and theta value for Dargay's model
3.	*urbanization (φ)*	Percentage point of urban citizens on total population	World bank's World Development indicators database from 2003 – 2015	Used for re-estimate beta and theta value for Dargay's model
4.	*Θ*	Speed of adjustment short-term toward long-term, $0 < \theta < 1$	Dargay, Gately and Sommer (2007)	Econometric estimation of parameters using annual data over the period 1960 – 2002
5.	*Γ*	Saturation level of vehicles per capita	Dargay, Gately and Sommer (2007)	Econometrical estimation of parameters using annual data over the period 1960 – 2002
6.	*GDP*	Real GDP per capita measured at purchasing power parity (PPP), (GDP per capita expressed) in 2010 USD (thousands), PPPs	OECD, World bank re-port. The data is collected from 2003 – 2015 for evaluation and projection	GDP data for some countries which not available in OECD report are acquired from World bank data with PPPs adjusted and 2011 USD
7.	*A*	Common parameter defines the shape of the vlr function	Dargay, Gately and Sommer (2007)	Econometrical estimation of parameters using annual data over the period 1960 – 2002
8.	*B*	Specific parameter for each country defines the shape of the vlr function	Dargay, Gately and Sommer (2007)	Econometrical estimation of parameters using annual data over the period 1960 – 2002 as the starting point. Then adjusted the β so that the sum of total actual car sale is within ±2 per cent of predicted value for the period from 1997 – 2007
9.	v_{it}	Historical data of stock of vehicle from 2003 – 2015	The International Organization of Motor Vehicle Manufacturers (OICA); report from the European Automobile Manufacturers Association (ACEA)	Used for both projection and evaluation of the passenger car sales model

Source: Own processing.