Forecasting German Car Sales Using Google Data and Multivariate Models

Fantazzini, Dean and Toktamysova, Zhamal

Moscow School of Economics, Moscow State University (Russia), Faculty of Economics, Higher School of Economics, Moscow (Russia)

2015

Forecasting German Car Sales Using Google Data and Multivariate Models

Dean Fantazzini* Zhamal Toktamysova †

Abstract

Long-term forecasts are of key importance for the car industry due to the lengthy period of time required for the development and production processes. With this in mind, this paper proposes new multivariate models to forecast monthly car sales data using economic variables and Google online search data. An out-of-sample forecasting comparison with forecast horizons up to 2 years ahead was implemented using the monthly sales of ten car brands in Germany for the period from 2001M1 to 2014M6. Models including Google search data statistically outperformed the competing models for most of the car brands and forecast horizons. These results also hold after several robustness checks which consider nonlinear models, different out-of-sample forecasts, directional accuracy, the variability of Google data and additional car brands.

Keywords: Car Sales, Forecasting, Google, Google Trends, Global Financial Crisis, Great Recession.

JEL classification: C22, C32, C52, C53, L62.

*Moscow School of Economics, Moscow State University, Leninskie Gory, 1, Building 61, 119992, Moscow, Russia. Fax: +7 4955105256 . Phone: +7 4955105267 . E-mail: fantazzini@mse-msu.ru .

†Faculty of Economics, Higher School of Economics, Moscow (Russia)

This is the working paper version of the paper *Forecasting German Car Sales Using Google Data and Multivariate Models*, forthcoming in the International Journal of Production Economics.

1 Introduction

Long-term forecasting of car sales plays an important role in the automobile industry. Accurate predictions allow firms to improve market performance, minimize profit losses, and plan manufacturing processes and marketing policies more efficiently.

Tough competition, significant investments, and the need for quick model updates are the specifics of the automotive industry which make forecasting an element of key importance for the sales and production processes. Like other complex industries, it can be characterized by long product development cycles varying from 12 up to 60 months. An effective planning of the production therefore requires accurate long-term sales forecasts. Inaccurate forecasts may result in several negative consequences, such as overstocking or shortage of production supplies, high costs for different workforce activities, loss of reputation for the manufacturer and even bankruptcy.

There are several economic factors affecting the automobile industry, and they can be broadly divided into three groups. The first group incorporates the technological aspects of the products: quality, innovation and technology, performance and economy of the engine, functionality, safety, space management, design and aesthetics (Lin and Zhang, 2004; Sa-ngasoongsong and Bukkapatnam, 2011). The second group comprises promotion and sales factors, including wholesale and retail prices, customer service, advertising campaigns, and brand image (Landwehr, Labroo, and Herrmann, 2011). These factors are significant, but usually do not have a long-term effect and automobile producers in most cases can manage and control them (Dekimpe, Hanssens, and Silva-Risso, 1998; Nijs, Dekimpe, Steenkamp, and Hanssens, 2001; Pauwels, Hanssens, and Siddarth, 2002; Pauwels, Silva-Risso, Srinivasan, and Hanssens, 2004). The third group includes various political, economic and social environmental factors which are generally beyond the control of manufacturers, such as organizational issues, political issues, global economic growth, ecological and physical forces, socio-cultural effects and consumer behavior. The use of these factors for car sales forecasting has been rather limited, see Brühl, Borscheid, Friedrich, and Reith (2009), Shahabuddin (2009), Wang, Chang, and Tzeng (2011) and Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012). Moreover, most previous studies have focused on the dynamics of car sales in the short-term, with forecast horizons usually less than 4 months, whereas car sales forecasting requires time scales with duration up to one year or more.

Following the growing number of Internet users (International Telecommunications Union, 2014) and the increasing popularity of Google as a search engine for obtaining information about cars, we propose the use of Google search data as a leading indicator for the long-term forecasting of car sales. In this regard, Google Search holds the world leadership among all search engines with a 54% market share (Net Applications, 2014). Since 2004, it has offered a tool called Google Trends, which provides information on the relative interest of users in a particular search query, at a given geographic region and at a given time (the data are available on a weekly or even daily basis). Moreover, Google Trends can attribute queries to different search categories (Autos, Computers, Finance, Health and others). In recent years, researchers worldwide have begun to use online search data to produce real-time forecasts where information from official sources is released with a lag (such as 'nowcasting'), or simply as an additional variable for forecasting purposes, see Choi and Varian (2012), Askitas and Zimmermann (2009), Suhoy (2009), Ginsberg, Mohebbi, Patel, Brammer, Smolinski, and Brilliant (2009), Da, Engelberg, and Pengjie (2011), D'Amuri and Marcucci (2013) and Fantazzini and Fomichev (2014) for some recent applications.

With this in mind, we propose a set of models for the long-term forecasting of car sales in Germany, which consider both economic variables and online search queries. Germany is the third biggest car producer in the world (about 14 million vehicles in 2013 and 20% of the total world production) and the absolute leader in Europe (31% of the total European production), see the reports by the German Association of the Automotive Industry (GTAI, 2014) and the Germany Trade and Invest Organization (VDA, 2014) for more details. As for Internet users, Germany has the second highest number of users in Europe (12.3% of all European users) and the 7th in the world. In June 2014, more than 71 million people in Germany visited the Web at least once a month, representing 88.6% of the adult population (Internet World Stats, 2014).

The first contribution of this paper is a set of multivariate models which include both Google data and economic variables. So far, the vast majority of the literature has used Google data as an exogenous variable in univariate models for short-term forecasting. Given that the car industry is interested in long-term forecasting, simple univariate models are not sufficient, and multivariate models are required to produce multi-step ahead forecasts for all variables, Google data included. Moreover, we consider multivariate models for both deseasonalized data, the usual approach in the economic literature, and for data not seasonally adjusted, which is more common in practice, since planning and production departments tend to work with raw data[1].

The second contribution of our paper is a large-scale forecasting exercise for ten car brands in Germany, where we compute out-of-sample forecasts ranging from 1 month to 24 months ahead. Our results show that models including car sales, Google data and economic variables outperform the competing models in the medium term for most of the car brands, while multivariate models including only car sales and Google data outperform the other models for long-term forecasts up to 24 steps ahead. The use of parsimonious models is crucial to obtain precise forecasts in the long run, and the use of Google search data represents a simple and powerful way to summarize the large amount of information available (see also Fantazzini and Fomichev, 2014).

The third contribution of the paper is a set of robustness checks to verify that our results also hold when considering nonlinear models, different out-of-sample forecasts, the use of directional accuracy as the main evaluation tool, Google data downloaded on different days, and additional car brands.

The paper is organized as follows. Section 2 describes the data and the in-sample analysis, and the forecasting models and their out-of-sample performance are reported in Section 3. Robustness checks are discussed in Section 4, and Section 5 briefly concludes.

2 Data and In-Sample analysis

We analyze new car registrations in the Federal Republic of Germany, as provided in press releases by the Federal Motor Transport Authority (Kraftfahrt-Bundesamt). These data cover the period from January 2001 to June 2014, for a total of 162 observations. The data consist of monthly numbers of new vehicle registrations by vehicle type and new registrations of passenger cars by brand starting from 2001. For different reasons, the information for some car brands was truncated: certain brands were present only after 2001; others stopped being observed well before 2014; or the registration statistics were not published due to the small number of registrations per month. Our car brands were selected based on the availability of a long time series for new car registrations and their presence in the "`Vehicle Brands`" Google subcategory. Moreover, car brands were chosen to reflect both foreign and domestic car producers.

There were only 22 brands which had both monthly data continuously available since 2001 and were present in Google Trends. We divided these brands into clusters by taking the average sales for each brand and using the method of k-means with Euclidian distance. We wanted to determine large, medium and small car manufacturers, and assign all brands into three clusters. The method of k-means allowed us to define the number of clusters a priori and minimize the within-cluster distance while maximizing the between-cluster distance (see e.g. Hartigan (1975)). The initial k cluster centers are chosen to maximize the initial distance. The data are arranged to the nearest cluster center, therefore k clusters are formed. Next, new cluster centers are chosen as centers of mass for the clusters. After recalculation, the data are again assigned to the nearest cluster centers. The procedure ends when all centers of mass are stabilized. We found three clusters consisting of the following brands:

- Large sellers: Volkswagen, Opel, Ford, BMW, Audi (average monthly sales between 19523 and 53820);
- Medium-sized sellers: Renault, Toyota, Peugeot, Hyundai, Fiat, Mazda, Citroen, Nissan (average monthly sales between 4976 and 14074);
- Small sellers: Jaguar, Kia, Land Rover, Porsche, Subaru, Honda, Volvo, Mitsubishi, Suzuki (average monthly sales between 355 and 3351).

We also used the method of k-means with the monthly sales data from January 2001 to June 2014 and we obtained the same division into three clusters.

For the sake of space, interest and to keep the empirical analysis computationally tractable, throughout the paper we will consider three large sellers (Volkswagen, Opel, BMW), three medium-sized sellers (Toyota, Fiat, Citroen), and four small sellers (Jaguar, Kia, Mitsubishi, Suzuki). The remaining 12 brands will be examined as a robustness check in section 4.5.

[1]The authors wish to thank an anonymous director of marketing and sales for pointing out this issue.

The plots of the monthly sales are reported in Figure 1 (right vertical axis). Car sales are subject to seasonal fluctuations and all car brands tend to show several peaks during the year, with the biggest one taking place at the end of spring. In general, car sales decline during winter. The Census X-12 tests for seasonality detected that all brands exhibit stable seasonality, with no evidence of moving seasonality.

The second source of data consists of Google Trends data, which can be downloaded from `www.google.com/trends/`, using the specific "`Autos and Vehicles`" category and its "`Vehicle Brands`" subcategory. The Google Index (GI) is the ratio of the number of queries relative to a particular category (in our case the car brand), with respect to all queries in the selected region at a given point of time. The data were collected for the whole of Germany for the period January 2004 - June 2014. The data have a weekly frequency and were converted to a monthly series by taking average values. While the GIs for a keyword are normalized to be bounded between 0 to 100, where 100 is the peak of the search queries, the GIs for a category are expressed in terms of percentage change from their first observation in January 2004, so that they can be both positive and negative. Their plots are reported in Figure 1 (left vertical axis): it is interesting to note that the turning points in the GIs anticipate those in the car sales for all car brands. This initial evidence suggests that Google data may be of some help for medium- and long-term forecasting.

Additionally, we included a number of economic variables related to car sales, based on recent works by Shahabuddin (2009) and Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012). These variables are assumed to reflect the state of the national economy, and the factors that can influence a consumer's decision to purchase a car. The selected economic variables and their descriptions are presented in Table 1. The data were collected for the period January 2001 to June 2014. All data, with the exception of building construction orders (which were available only seasonally adjusted), show some form of seasonality, with peaks during the summer season and troughs at the end of the year. The quarterly GDP data were converted to monthly data via the quadratic match average procedure, while the daily data for Euribor rates were transformed into monthly data by taking their average. Their plots are reported in Figure 2.

Economic variable	**Frequency**	**Seasonally adjusted**	**Source**	**Explanation**
Building Construction (BC)	M	yes	GFB	Volume index of new orders for residential buildings construction
Consumer Confidence Indicator (CCI)	M	no	DG ECFIN	Consumer survey that reflects consumer expectations
Consumer Price Index (CPI)	M	no	FSO	Measure of the ratio of a price of fixed set of consumer goods and services in current period to its price in a basic period
Euro Interbank Offered Rate (EURIBOR)	D	no	EBF	Calculated as an average rate of lending rate of the banks which participate in the survey. For the current research EURIBOR for long-term credits (1 year) is considered
Gross Domestic Product (GDP)	Q	no	FSO	Market value of all goods and services produced within a country. In the present work GDP in nominal billions Euro was taken
Production Index (PI)	M	no	FSO	Production Index for durable goods
Unemployment Rate (UR)	M	no	FEA	The registered unemployed population as a percentage of the civilian labor force
Petrol Price (PP)	M	no	FSO	Consumer price for petrol, price index

Table 1: Description of economic variables used in the analysis. The second column reports the frequency of publishing: M - monthly data, Q - quarterly data, D - daily data. The abbreviations used in the fourth column represent the data sources: GFB - German Federal Bank (Deutsche Bundesbank), DG ECFIN -Directorate General for Economic and Financial Affairs, FSO - The Federal Statistical Office (Statistisches Bundesamt), EBF - The European Banking Federation, FEA - The Federal Employment Agency (Bundesagentur für Arbeit).

Data with seasonal behavior were seasonally adjusted with the Census X-12 adjustment program developed by US Census Bureau. However, we also considered the raw data, since they are more common in practice and of greater interest for production planners and marketing managers, who base their decisions on real data which exhibit seasonality.

All data were transformed into logarithms to reduce variability and convert nonlinear patterns to

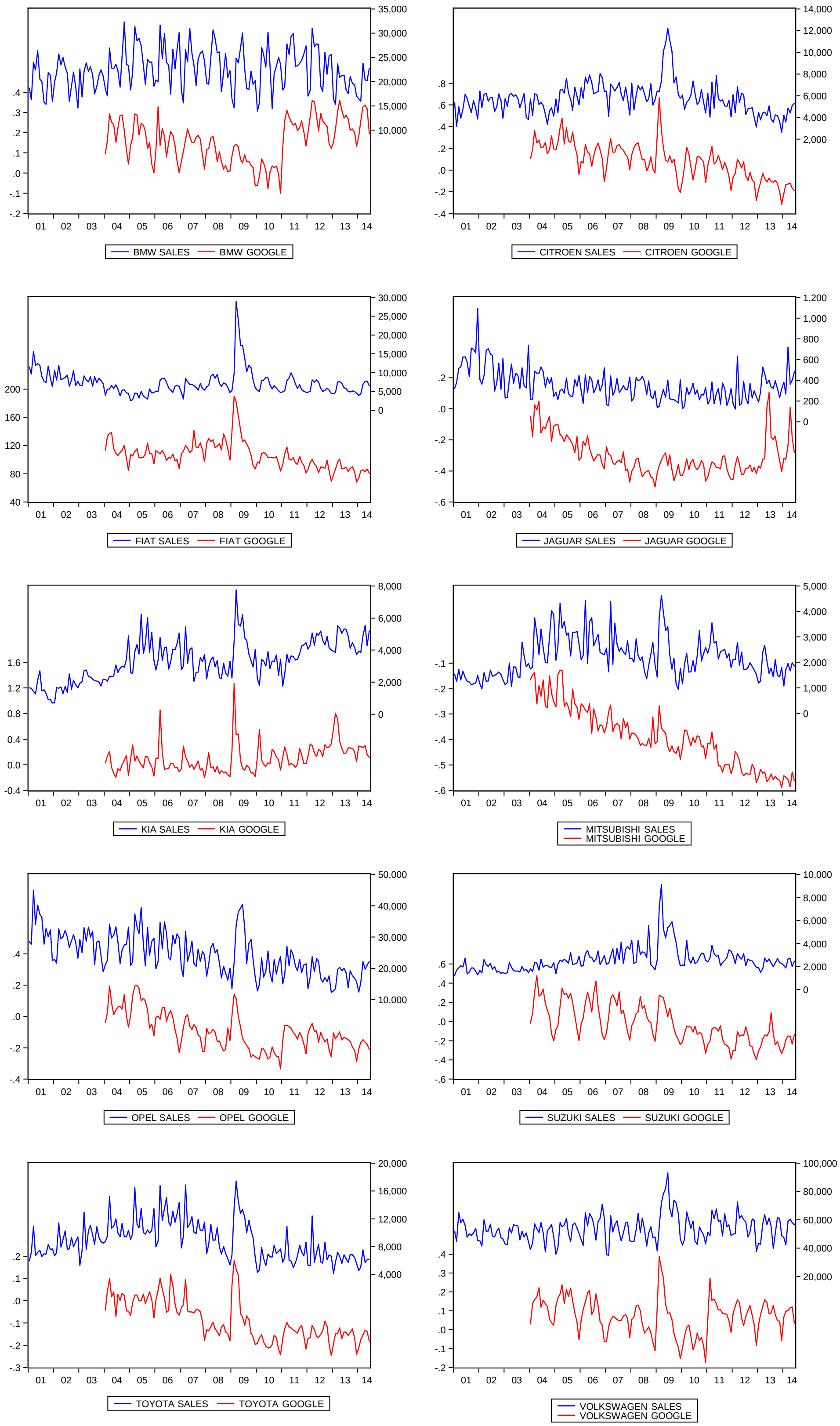

Figure 1: Car sales (right vertical axis) and relative GIs (left vertical axis) - not seasonally adjusted. Sample: 2001M1 - 2014M6.

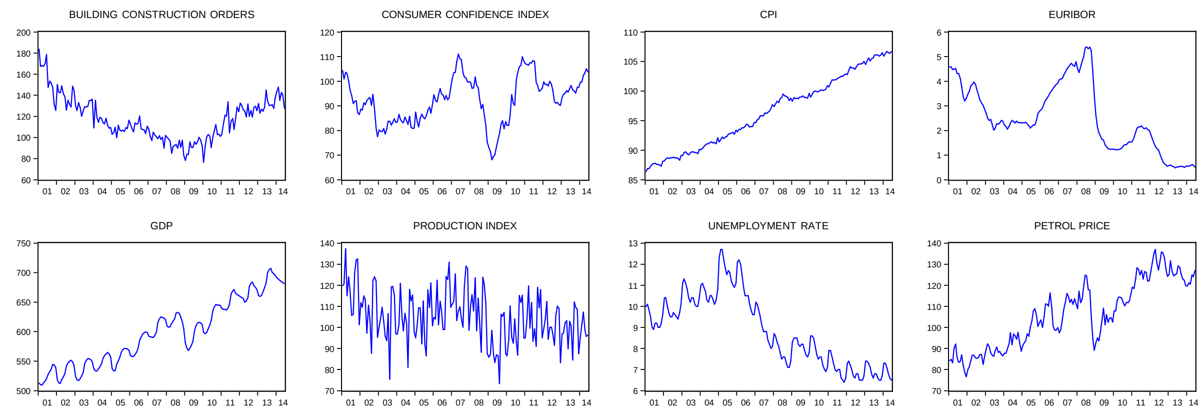

Figure 2: Economic variables - not seasonally adjusted. Sample: 2001M1 - 2014M6

linear patterns[2](see Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012)). The descriptive statistics for the car registrations, the Google data and the economic variables (both seasonally adjusted and raw data) are not reported for the sake of space and are available from the authors upon request.

To select the best multivariate model for each car brand, we follow the structural relationship identification methodology discussed by Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012) for the case of the US car market. Briefly, the first step is to identify the order of integration using unit root tests; if all variables are stationary, VAR and VARX (Vector Autoregressive with exogenous variables) models are used. The second step determines the exogeneity of each variable using the sequential reduction method for weak exogeneity by Hall, Henry, and Greenslade (2002), who consider weakly exogenous each variable for which the test is not rejected and re-test the remaining variables until all weakly exogenous variables are identified. For non-stationary variables, cointegration rank tests are employed to determine the presence of a long-run relationship among the endogenous variables: if this is the case, VECM or VECMX (Vector Error Correction model with exogenous variables) models are used, otherwise VAR or VARX models in differences are applied. The last step is to compute the impulse response functions from the chosen model to trace the effect of a unit shock in one of the variables on the future values of car sales, and to compute out-of-sample forecasts (see Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012) for more details). Our approach differs from the one proposed by Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012) in two respects: first, we employ unit root tests and cointegration tests allowing for structural breaks, given the possible break in the years 2008-2009 during the global financial crisis. Second, we employ the previous identification methodology for both the seasonally adjusted data and the raw data.

2.1 Stationarity

2.1.1 Seasonally Adjusted data

The stationarity of our variables is analyzed using several unit root tests allowing for potential endogenous structural break(s), both under the null of a unit root and under the alternative. We justify this choice considering the strong influence the global financial crisis in the years 2007-2009 had on the German economy, which is visible when looking at Figures 1 and 2. As for the Google data, we remark that the statistical effects of dividing the original search data by the total number of web searches in the same week and area are unknown, so that we cannot say a priori whether they are stationary or not (see also Fantazzini and Fomichev (2014) for a discussion on this issue). More specifically, we employed four unit root tests: the Lee and Strazicich (2003) unit root tests allowing for one and two breaks, respectively, and the Range Unit Root (RUR) and the Forward-Backward RUR tests suggested by Aparicio, Escribano, and Garcia (2006), which are non-parametric tests robust against nonlinearities, error distributions, structural breaks and outliers. A brief description of these tests is reported in the Technical Appendix

[2]The GIs were linearly re-scaled to positive numbers and then transformed into logarithms.

A accompanying this paper and can be found on the authors' websites.

	RUR *Test statistic*	**FB** *Test statistic*	**LS 1 break** *Test statistic*	**LS 2 breaks** *Test statistic*	**The null hypothesis is rejected by all tests?**
			Car sales		
BMW	0.71 *	1.16	-5.08 *	-11.14 *	no
Citroen	1.34	1.95	-5.12 *	-6.09 *	no
Fiat	0.79 *	1.89	-4.75 *	-6.31 *	no
Jaguar	0.87 *	1.39	-4.47	-6.98 *	no
Kia	1.42	2.01	-4.94 *	-5.89 *	no
Mitsubishi	0.79 *	1.34	-5.05 *	-5.79 *	no
Opel	0.87 *	1.56	-6.17 *	-6.87 *	no
Suzuki	1.02 *	1.67	-4.91 *	-6.47 *	no
Toyota	1.50	1.95	-4.92 *	-5.86 *	no
Volkswagen	0.87 *	1.73	-6.66 *	-7.52 *	no
			Economic variables		
BUILD	1.34	2.17	-2.33	-8.68 *	no
CCI	1.18	2.23	-3.60	-4.07	no
CPI	9.14 *	13.15*	-3.53	-4.10	no
EURIBOR	3.07	3.73 *	-3.46	-4.29	no
PP	2.68	3.96 *	-3.65	-5.26	no
GDP	6.30 *	8.75 *	-3.67	-4.53	no
PI	1.42	1.67	-3.88	-4.80	no
UR	5.28 *	7.30 *	-3.42	-5.66	no
			Google data		
BMW GI	1.34	1.77	-5.24 *	-8.59 *	no
Citroen GI	1.97	2.34	-5.98 *	-6.71 *	no
Fiat GI	1.43	2.34	-4.59 *	-7.07 *	no
Jaguar GI	1.52	1.90	-7.12 *	-8.10 *	no
Kia GI	0.80 *	1.39	-7.45 *	-8.12 *	no
Mitsubishi GI	2.68	2.97	-9.26 *	-9.83 *	no
Opel GI	1.25	2.53	-4.51 *	-5.24	no
Suzuki GI	1.88	2.09	-7.18 *	-8.24 *	no
Toyota GI	1.34	1.90	-4.67 *	-5.17	no
Volkswagen GI	1.34	1.83	-4.96 *	-5.55	no

Table 2: Unit root tests: RUR = Range Unit Root test by Aparicio, Escribano, and Garcia (2006); FB = Forward-Backward RUR test by Aparicio, Escribano, and Garcia (2006); LS = Unit Root test by Lee and Strazicich (2003). Null hypothesis: the time series has a unit root. * Significance at the 5% level.

The results in Table 2 show that the majority of our time series are not stationary. However, the Lee and Strazicich (2003) tests show a stronger evidence of unit roots for economic variables, while the Aparicio, Escribano, and Garcia (2006) tests show the same for car sales and Google data. If we follow a conservative approach and analyze when all four tests reject the null hypothesis (see the last column in Table 2), then all car brands can be deemed non-stationary.

2.1.2 Raw data

To test the null hypothesis of a periodic unit root, we follow the two-step strategy suggested by Boswijk and Franses (1996) and Franses and Paap (2004). In the first step, a likelihood ratio test for testing a single unit root in a Periodic Auto-Regressive (PAR) model of order p is performed. Since there is no version of this test with endogenous breaks, we estimated it both with the full sample starting in 2001, and with a smaller sample starting in 2008. The year 2008 was chosen following the previous evidence of a possible break in this year, which emerged with the unit root tests allowing for breaks in the case of seasonally adjusted data. If the null of a periodic unit root cannot be rejected, Boswijk and Franses (1996) and Franses and Paap (2004) suggest to test in a second step whether the process contains a non-periodic unit root equal to 1 for all seasons. A description of these tests is reported in the Technical Appendix B.

Table 3 shows that car sales offer different results depending on the sample used: if the full sample is considered, non-stationarity is rejected for all car brands but BMW (for which the estimates did not reach numerical convergence); if the smaller sample starting from 2008 is used, the test failed to converge for several brands, while for two brands (Citroen and Kia) the null of a non-periodic unit root cannot be rejected. This evidence again highlights the possible presence of a structural break in 2008 during the global financial crisis. Economic variables and GIs are mostly non-stationary with a non-periodic unit root and the results do not change substantially with the sample used.

	Sample: 2001-2014		Sample: 2008-2014	
	1st step	**2nd step**	**1st step**	**2nd step**
	H_0: **periodic unit root**	H_0: **non periodic unit root**	H_0: **periodic unit root**	H_0: **non periodic unit root**
		Car Sales		
BMW	NC	NC	NC	NC
Citroen	18.66*	/	7.21	0.46
Fiat	16.60*	/	4.43	**0.00**
Jaguar	42.41*	/	NC	NC
Kia	10.46*	/	4.96	0.08
Mitsubishi	22.97*	/	16.96*	/
Opel	15.38*	/	10.66*	/
Suzuki	24.85*	/	15.95*	/
Toyota	10.19*	/	15.81*	/
Volkswagen	58.20*	/	NC	NC
		Economic Variables		
BUILD	7.99	0.09	2.32	0.11
CCI	3.23	0.06	1.02	0.14
CPI	0.13	**0.00**	0.30	0.44
EURIBOR	0.37	0.66	1.99	0.15
PP	1.97	0.88	1.36	0.10
GDP	0.01	**0.00**	0.15	**0.00**
PI	36.79*	/	22.07*	/
UR	0.52	0.56	NC	NC
		Google data		
BMW GI	8.93	0.49	2.71	0.53
Citroen GI	4.90	0.47	4.46	0.13
Fiat GI	4.47	**0.04**	1.84	0.11
Jaguar GI	12.02*	/	5.17	**0.01**
Kia GI	16.82*	/	8.07	0.76
Mitsubishi GI	3.91	0.99	2.19	0.35
Opel GI	6.06	0.64	6.69	0.53
Suzuki GI	3.60	**0.02**	3.63	**0.04**
Toyota GI	5.86	0.46	5.15	**0.01**
Volkswagen GI	11.20*	/	5.38	0.39

Table 3: Periodic Unit root tests by Boswijk and Franses (1996) and Franses and Paap (2004).
* Significance at the 5% level. NC = Not Converged. The second step is performed only if the first step numerically converged and did not reject the null hypothesis. p-values smaller than 0.05 are in bold.

2.2 Weak Exogeneity and Cointegration Tests

2.2.1 Seasonally Adjusted data

The next step in the structural relationship identification methodology discussed by Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012) is to determine the exogeneity of each variable using the sequential reduction method for weak exogeneity proposed by Hall, Henry, and Greenslade (2002). This method exogenizes all weakly exogenous variables and re-tests the remaining variables until all weakly exogenous variables are identified. The variables that reject the null of weak exogeneity after re-testing are reported in Table 12 in Appendix A: the Euribor series can be considered weakly exogenous for four car brands, while almost all other variables are deemed endogenous (with some exceptions for Mitsubishi).

We then proceeded to test for cointegration using the variables which were deemed endogenous according to the previous sequential test procedure by Hall, Henry, and Greenslade (2002). We test for cointegration using a set of cointegration tests allowing for the presence of structural break(s):

- Gregory and Hansen (1996) single-equation cointegration test allowing for one endogenous break;
- Hatemi (2008) single-equation cointegration test allowing for two endogenous breaks;
- Johansen, Mosconi, and Nielsen (2000) multivariate test allowing for the presence of one or two exogenous break(s), where the dates of the breaks are the ones selected by the Gregory and Hansen (1996) and Hatemi (2008) tests, respectively.

A description of these cointegration tests is reported in the Technical Appendix C. For the sake of generality, we also considered the multivariate cointegration test by Johansen (1995) without breaks. The main advantage of single-equation approaches is that they allow for endogenous breaks. However, these tests are not suitable when the right-hand variables in the cointegration vector are not weakly exogenous (as in our case) and when there is more than one cointegrating vector. In this case, multivariate cointegration tests should be used. The only problem with the multivariate tests by Johansen, Mosconi, and Nielsen (2000) is that they allow only for exogenous breaks. Accordingly, we followed a 2-step strategy: we first estimated the single-equation tests to obtain an indication of the structural break dates. We then used these dates to compute the tests by Johansen, Mosconi, and Nielsen (2000). Finally, we remark that the number of lags for the Johansen tests were chosen to minimize the Schwartz criterion and to make the residuals approximately white noise.

Single-Equation cointegration tests				
	Gregory and Hansen (1996) one(endogenous) break		**Hatemi (2008) two(endogenous) breaks**	
	Z-t statistic	*Break date*	*Z-t statistic*	*Break dates*
BMW	-10.61*	2010M02	-11.14*	2006M09 2008M07
Citroen	-7.38*	2009M02	-8.35	2005M08 2007M07
Fiat	-7.54*	2006M01	-8.27	2005M11 2007M08
Jaguar	-14.54*	2012M09	-14.30*	2007M10 2011M02
Kia	-8.27*	2006M09	-8.61	2006M09 2011M01
Mitsubishi	-10.98*	2009M03	-10.79*	2008M04 2008M12
Opel	-8.72*	2009M02	-7.60	2009M09 2010M10
Suzuki	-10.85*	2009M02	-10.14	2006M09 2007M06
Toyota	-7.95*	2009M12	-8.40	2006M09 2009M07
Volkswagen	-9.96*	2009M03	-9.35	2005M08 2007M08

Multivariate cointegration tests					
	Johansen (1995) No Breaks	**Johansen, Mosconi, and Nielsen (2000) one(exogenous) break**		**Johansen, Mosconi, and Nielsen (2000) two (exogenous) breaks**	
	N. of CEs at 5% level	*N. of CEs at 5% level*	*Break date (GH,1996)*	*N. of CEs at 5% level*	*Break dates (H,2008)*
BMW	5 CE	5 CE	2010M02	5 CE	2006M09 2008M07
Citroen	5 CE	4 CE	2009M02	5 CE	2005M08 2007M07
Fiat	7 CE	5 CE	2006M01	7 CE	2005M11 2007M08
Jaguar	5 CE	4 CE	2012M09	5 CE	2007M10 2011M02
Kia	5 CE	3 CE	2006M09	4 CE	2006M09 2011M01
Mitsubishi	4 CE	0 CE	2009M03	*NC*	2008M04 2008M12
Opel	5 CE	4 CE	2009M02	5 CE	2009M09 2010M10
Suzuki	5 CE	5 CE	2009M02	*NC*	2006M09 2007M06
Toyota	5 CE	5 CE	2009M12	5 CE	2006M09 2009M07
Volkswagen	5 CE	5 CE	2009M03	5 CE	2005M08 2007M08

Table 4: Single-equation and multivariate cointegration tests with and without structural break(s) for seasonally-adjusted data. The null hypothesis for all tests is the absence of cointegration. The tests considered the case of a level shift. The table cells for the Johansen tests report the number of CEs selected at the 5% level. NC=not converged. * Significance at the 5% level.

Table 4 shows that there is strong evidence for cointegration for all considered car brands. However, structural breaks seem to have a non-negligible effect, particularly when considering Johansen multivariate tests. Moreover, the effects of breaks appear to be much stronger for foreign brands than for domestic brands (BMW, Volkswagen and, to a lesser extent, Opel), for which the cointegration tests do not change substantially when breaks are taken into account.

2.2.2 Raw data

To determine the exogeneity of variables with potential seasonal behavior, we extend the previous sequential reduction method for weak exogeneity by including centered seasonal dummies: they sum to zero over time and therefore do not affect the asymptotic distributions of the tests (see Johansen (1995, 2006)). The variables that reject the null of weak exogeneity after re-testing are reported in Table 13 in Appendix A: the results for raw data are not too dissimilar to the seasonally-adjusted data, even though there are less variables which are weakly exogenous. We then tested for cointegration using the variables which were found to be endogenous, and the previous cointegration tests augmented with centered seasonal dummies, see Table 5.

In the case of raw data, the evidence for cointegration appears to be quite similar to that of seasonally-adjusted data, particularly when considering the Johansen test without breaks and with one break. Moreover, the fact that the Johansen test with two breaks failed to converge for some car brands indicates that our sample is too small for two breaks and that only tests with one break should be considered.

Periodic cointegration tests using all variables could not be implemented due to the high number of parameters being estimated (the so-called "curse of dimensionality"). However, we wanted to consider a restricted bivariate periodic error correction model including only car sales and Google data. Even though such a specification is definitely biased – missing several important economic variables – this

	Single-Equation cointegration tests			
	Gregory and Hansen (1996) one (endogenous) break		**Hatemi (2008) two (endogenous) breaks**	
	Z-t statistic	*Break date*	*Z-t statistic*	*Break dates*
BMW	-10.78*	2010M02	11.35*	2006M09 2008M07
Citroen	-7.70*	2009M02	8.60	2005M08 2007M07
Fiat	-7.63*	2005M10	8.64	2005M10 2007M08
Jaguar	-13.10*	2006M11	NC	NC
Kia	-8.71*	2006M09	9.25	2009M09 2011M01
Mitsubishi	-11.54*	2009M02	10.88*	2008M03 2008M12
Opel	-8.48*	2009M02	7.30	2009M09 2010M12
Suzuki	-11.00*	2009M02	9.64	2006M09 2007M07
Toyota	-7.44*	2009M12	8.03	2009M10 2010M12
Volkswagen	-10.67*	2009M02	9.63	2005M08 2007M07

	Multivariate cointegration tests				
	Johansen (1995) No Breaks	**Johansen, Mosconi, and Nielsen (2000) one(exogenous) break**		**Johansen, Mosconi, and Nielsen (2000) two(exogenous) breaks**	
	N. of CEs at 5% level	*N. of CEs at 5% level*	*Break date (GH,1996)*	*N. of CEs at 5% level*	*Break dates (H,2008)*
BMW	5 CE	4 CE	2010M02	5 CE	2006M09 2008M07
Citroen	5 CE	5 CE	2009M02	5 CE	2005M08 2007M07
Fiat	5 CE	6 CE	2005M10	7 CE	2005M10 2007M08
Jaguar	3 CE	0 CE	2006M11	NC	NC
Kia	5 CE	5 CE	2006M09	5 CE	2009M09 2011M01
Mitsubishi	4 CE	4 CE	2009M02	NC	NC
Opel	5 CE	4 CE	2009M02	5 CE	2009M09 2010M12
Suzuki	5 CE	6 CE	2009M02	NC	NC
Toyota	5 CE	5 CE	2009M12	5 CE	2009M10 2010M12
Volkswagen	5 CE	6 CE	2009M02	6 CE	2005M08 2007M07

Table 5: Single-equation and multivariate cointegration tests with and without structural break(s) for raw data. The null hypothesis for all tests is the absence of cointegration. The tests considered the case of a level shift. The table cells for the Johansen tests report the number of CEs selected at the 5% level. NC=not converged. * Significance at the 5% level.

parsimonious model can nevertheless be of interest for forecasting purposes. Moreover, the capacity of Google data to summarize a wealth of information should not be underestimated. In this regard, we implemented the single-equation periodic cointegration test discussed in Franses and Paap (2004), which is an extension of the Boswijk (1994) cointegration test. The null hypothesis is the absence of cointegration against the alternative of periodic cointegration and the right-hand variables should be weakly exogenous. A description of this test as well as the test for weak exogeneity in the case of periodic variables by Boswijk (1994) is reported in the Technical Appendix D. Since we are not aware of any extension of this test allowing for structural breaks, we estimated it using both the full sample and a reduced sample starting in 2008 to take any potential break into account and the results are reported in Table 14 in Appendix A: the evidence in favor of periodic cointegration is fairly strong, but the results of the Boskwijk test statistics change partially when the smaller sample starting in 2008 is considered. Caution should therefore be exercised when dealing with this restricted model. Interestingly, the GIs are weakly exogenous with respect to car sales for almost all brands at the 5% level and this outcome does not change substantially with the sample used.

2.3 Impulse Response Functions

After the VECM (or VECMX) models were selected for each car brand, we proceeded to compute the impulse response functions (IRFs) in order to trace the effects of a one-time shock in one of the variables on current and future values of car sales. More specifically, we computed the generalized impulse response functions by Pesaran and Shin (1998), which do not depend on the ordering of the variables. For the sake of interest and space, we report here only the IRFs for the seasonally-adjusted sales data (Figure 3) with respect to a generalized one standard deviation innovation in the Google Indexes. Moreover, we report in Table 6 the estimated long-run parameters in the cointegration equations and their adjustment coefficients for the Volkswagen car sales equation, noting that Volkswagen is the biggest car maker and seller in Germany. A battery of misspecification tests computed on the VECMX model residuals is reported in the same table as well: we computed multivariate LM test statistics for residual serial correlation up to a specified order, univariate and multivariate Jarque-Bera residual normality tests, and the multivariate White heteroskedasticity test (see Johansen (1995) and Lutkepohl (2005) for more details about these tests). The full results are available from the authors upon request.

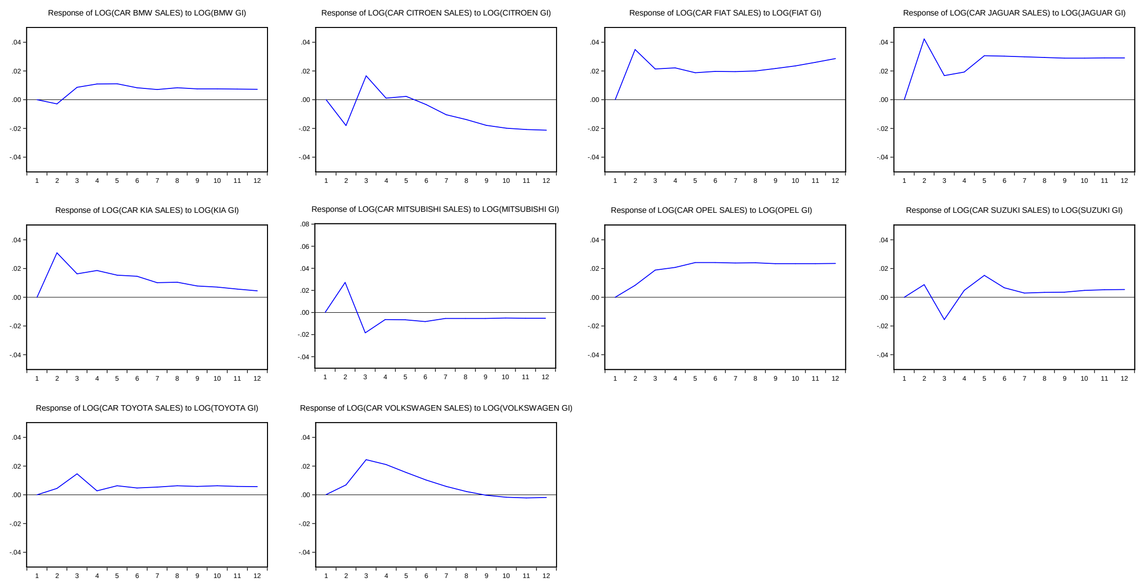

Figure 3: Impulse response functions: response of car sales (in logs) to generalized one standard deviation innovations in the Google Indexes.

	Long-run parameters (β)				
	CE 1	**CE 2**	**CE 3**	**CE 4**	**CE 5**
Log(SALES(-1))	1	0	0	0	0
Log(BC(-1))	0	1	0	0	0
Log(CCI(-1))	0	0	1	0	0
Log(CPI(-1))	0	0	0	1	0
Log(EURIBOR(-1))	0	0	0	0	1
Log(PP(-1))	0.11	-0.71	-0.29	-0.03	0.77
	[0.35]	[-2.63]	[-0.91]	[-3.34]	[1.20]
Log(PI(-1))	2.03	1.90	1.57	0.14	-8.97
	[5.19]	[5.60]	[3.97]	[14.10]	[-11.10]
Log(UR(-1))	0.98	-1.12	-0.27	0.05	6.53
	[3.51]	[-4.63]	[-0.95]	[6.33]	[11.31]
Log(GOOGLE(-1))	-1.95	-0.77	0.08	-0.03	-3.16
	[-6.82]	[-3.08]	[0.26]	[-3.83]	[-5.34]
Log(GDP(-1))	2.16	-2.96	-0.83	-0.46	19.78
	[2.91]	[-4.61]	[-1.11]	[-24.16]	[12.95]
Constant	-27.59	14.91	-4.73	-2.14	-89.07
	[-6.64]	[4.13]	[-1.13]	[-20.05]	[-10.39]
Adjustment coefficients (α) - car sales equation					
	-0.72	-0.04	0.64	5.55	0.18
	[-6.33]	[-0.23]	[4.60]	[1.70]	[4.25]

Misspecification tests			
	p-values		*p-values*
Multi. LM(1)	0.06	*Uni. JB test*	
Multi. LM(2)	0.76	SALES	**0.02**
Multi. LM(3)	0.22	BC	0.77
Multi. LM(4)	0.35	CCI	0.12
Multi. LM(5)	**0.02**	CPI	0.09
Multi. LM(6)	0.65	EURIBOR	0.62
Multi. LM(7)	0.75	PP	0.43
Multi. LM(8)	0.09	PI	**0.03**
Multi. LM(9)	0.52	UR	0.54
Multi. LM(10)	0.41	GOOGLE	**0.01**
Multi. LM(11)	0.06	GDP	0.51
Multi. LM(12)	0.33	*Multi. JB test*	**0.01**
		Multi. White	0.30

Table 6: Long-run parameters and adjustment coefficients for the Volkswagen car sales equation (left table). Misspecification tests on the residuals from the Volkswagen VECMX model (right table). t-statistics are reported in brackets, while p-values smaller than 5% are reported in bold.

As expected, a unit shock in the Google Index has a rather long and positive effect for almost all car brands. Similarly, the model estimates in Table 6 show that the Google Index enters almost all cointegration equations with significant positive coefficients[3], while the residual tests do not signal any serious misspecification.

[3]The signs of the long-run parameters in Table 6 are switched due to the error correction representation.

3 Out-of-Sample Forecasting Analysis

The last step in the structural relationship identification methodology discussed by Sa-ngasoongsong, Bukkapatnam, Kim, Iyer, and Suresh (2012) is to compare the forecasting performances of the selected VECM (or VECMX) models with a set of competitors.

3.1 Seasonally Adjusted data

We compared a set of 34 models, which allow for different degrees of model flexibility, parsimonious specifications and numerical tractability. More specifically, three types of multivariate models were employed:

- *Vector Error Correction (VEC) models:* We considered both VECM and VECMX models, as well as models with and without Google data, to better examine their effects on forecasting performance. The number of lags was selected to minimize the Schwartz criteria and to make the residuals approximately white noise. We also considered a set of parsimonious bivariate specifications including only car sales and Google data, which may be of interest for long-term forecasting.

- *Vector Auto-Regressive (VAR) models:* We considered VAR models with variables in log-levels and in log-differences, to consider both cases of stationarity and non-stationarity. Moreover, models with and without exogenous variables and with and without Google data were also considered. Finally, a set of parsimonious bivariate VAR models including only car sales and Google data was included.

- *Bayesian Vector Auto-Regressive (BVAR) models:* When there are a lot of variables and a high number of lags, estimating the parameters of a VAR model can be very difficult, if not impossible. One way to solve this issue is to shrink the parameters using Bayesian methods. Bayesian VAR models have recently enjoyed a lot of success in macroeconomic forecasting (see Koop and Korobilis (2010) for a recent review and Fantazzini and Fomichev (2014) for a recent application with Google data). In this regard, we used the so-called Litterman/Minnesota prior, which was developed by researchers at the University of Minnesota and at the Federal Reserve Bank of Minneapolis, and which is a common choice in empirical applications due to its computational speed and forecasting success (see Doan, Litterman, and Sims (1984), Litterman (1986) and Koop and Korobilis (2010)). A brief description of BVAR models can be found in the Technical Appendix E. Similarly to the VAR and VECM models, we considered models with and without exogenous variables, with and without Google data and with variables both in log-levels and in log-differences.

Besides these models, we also considered a set of standard univariate time series models:

- The Random Walk with drift;

- An AR(12) model for the log-returns of car sales.

Moreover, all models without Google data were estimated using both a long sample starting in 2001 and a short one starting in 2004, in the hope that this will show more clearly the advantages of Google data. The full details of all 34 multivariate models are reported in Table 7. For ease of reference, we also report in the sixth column a short-cut notation for identifying each model in the tables reporting the models forecasting performances.

We used the data between 2001M1 and 2008M9 as the first initialization sample for the models without Google data, and data from 2004M1 till 2008M9 for the models with Google data and those without Google data but estimated on a shorter sample. The evaluation period ranged from 2008M10 till 2014M6 and was used to compare forecasts from 1 step ahead up to 24 steps ahead. The top three models in terms of the Mean Squared Prediction Error (MSPE) for each forecasting horizon and each car brand are reported in Table 15, while the full results are available from the authors upon request.

Table 15 shows that there is no single model which outperforms all competitors for all horizons and all car brands. However, some general indications can be retrieved:

- The MSPEs of the competing models with forecasting horizons up to 8-10 steps ahead are relatively close (results not reported) and the Random Walk and the AR(12) models are sometimes ranked among the top three models;

Type	Log-levels / log-returns	Exogenous variables	Google data	Notes	Short cut notation (seas. adj. data)	Short cut notation (raw data)
				VEC MODELS		
VECM	Log-lev/log-ret	no	yes		VECM	VECMP
VECMX	Log-lev/log-ret	yes	yes		VECMX	VECMXP
VECM	Log-lev/log-ret	no	no		VECMNOGO	VECMPNOGO
VECM	Log-lev/log-ret	no	no	Sample starts in 2004	VECMNOGO4	VECMPNOGO4
VECMX	Log-lev/log-ret	yes	no		VECMXNOGO	VECMXPNOGO
VECMX	Log-lev/log-ret	yes	no	Sample starts in 2004	VECMXNOGO4	VECMXPNOGO4
VECM	Log-lev/log-ret	no	yes	Only sales and GI. Lags: 1,12	VECongo112	VEPongo112
VECM	Log-lev/log-ret	no	yes	Only sales and GI. Lags: 1-12	VECongo12	VECPongo12
				VAR MODELS		
VAR	Log-levels	no	yes		VAR	VARP
VAR	Log-returns	no	yes		VARD	VARPD
VAR	Log-levels	yes	yes		VARX	VARXP
VAR	Log-returns	yes	yes		VARXD	VARXPD
VAR	Log-levels	no	no		VARNOGO	VARPNOGO
VAR	Log-levels	no	no	Sample starts in 2004	VARNOGO4	VARPNOGO4
VAR	Log-returns	no	no		VARDNOGO	VARPDNOGO
VAR	Log-returns	no	no	Sample starts in 2004	VARDNOGO4	VARPDNOGO4
VAR	Log-levels	yes	no		VARXNOGO	VARXPNOGO
VAR	Log-levels	yes	no	Sample starts in 2004	VARXNOGO4	VARXPNOGO4
VAR	Log-returns	yes	no		VARXDNOGO	VARXPDNOGO
VAR	Log-returns	yes	no	Sample starts in 2004	VARXDNOGO4	VARXPDNOGO4
VAR	Log-levels	no	yes	Only sales and GI. Lags: 1,12	VARongo112	VARongo112
VAR	Log-levels	no	yes	Only sales and GI. Lags: 1-12	VARongo12	VARongo12
VAR	Log-returns	no	yes	Only sales and GI. Lags: 1,12	VADongo112	VADongo112
VAR	Log-returns	no	yes	Only sales and GI. Lags: 1-12	VADongo12	VADongo12
				BVAR MODELS		
BVAR	Log-levels	yes	yes		BVAR	BVARP
BVAR	Log-returns	yes	yes		BVARD	BVARPD
BVAR	Log-levels	yes	no		BVARNOGO	BVARPNOGO
BVAR	Log-levels	yes	no	Sample starts in 2004	BVARNOGO4	BVARPNOGO4
BVAR	Log-returns	yes	no		BVARDNOGO	BVARPDNOGO
BVAR	Log-returns	yes	no	Sample starts in 2004	BVARDNOGO4	BVARPDNOGO4
				UNIVARIATE TIME SERIES MODELS		
AR(12)	Log-returns	no	no		AR12	AR12
AR(12)	Log-returns	no	no	Sample starts in 2004	AR124	AR124
R. w.	Log-returns	no	no		RW	RW
R. w.	Log-returns	no	no	Sample starts in 2004	RW4	RW4
				PERIODIC ERROR CORRECTION MODELS		
Periodic ECM	Log-lev/log-ret	no	yes	Only sales and GI. Lags: 1-12	/	PECM

Table 7: Models used for forecasting (baseline case).

- Bayesian VAR models, particularly in differences and without Google data, perform rather well across all car brands and for short and medium forecasts (up to 12 steps ahead);
- Bivariate models including only car sales and Google models and using only the first and the 12th lags perform extremely well across most of the car brands examined, particularly for long-term forecasts. The parsimonious specifications of these models clearly allow for efficiency gains where forecasting is of concern.
- The forecasting power of the best models using Google data increases with the length of the forecast horizon, particularly with forecast horizons higher than 12 steps ahead. This evidence is similar to that found in D'Amuri and Marcucci (2013) and Fantazzini and Fomichev (2014).
- Models without Google data estimated with the long sample starting in 2001 tend to perform better than those estimated with a shorter sample starting in 2004.
- There are no particular differences between large, medium-sized and small sellers and between foreign and German manufacturers.

So as to provide an idea about how prediction errors evolve over time, Figure 4 (columns 1 and 2 for seasonally adjusted data) shows the ratios of the MSPE of the best model with Google data and the Random Walk model across all forecasting horizons, together with the ratios of the MSPE of the best model without Google data and the Random Walk model. We remark that the best models tend to vary across different horizons.

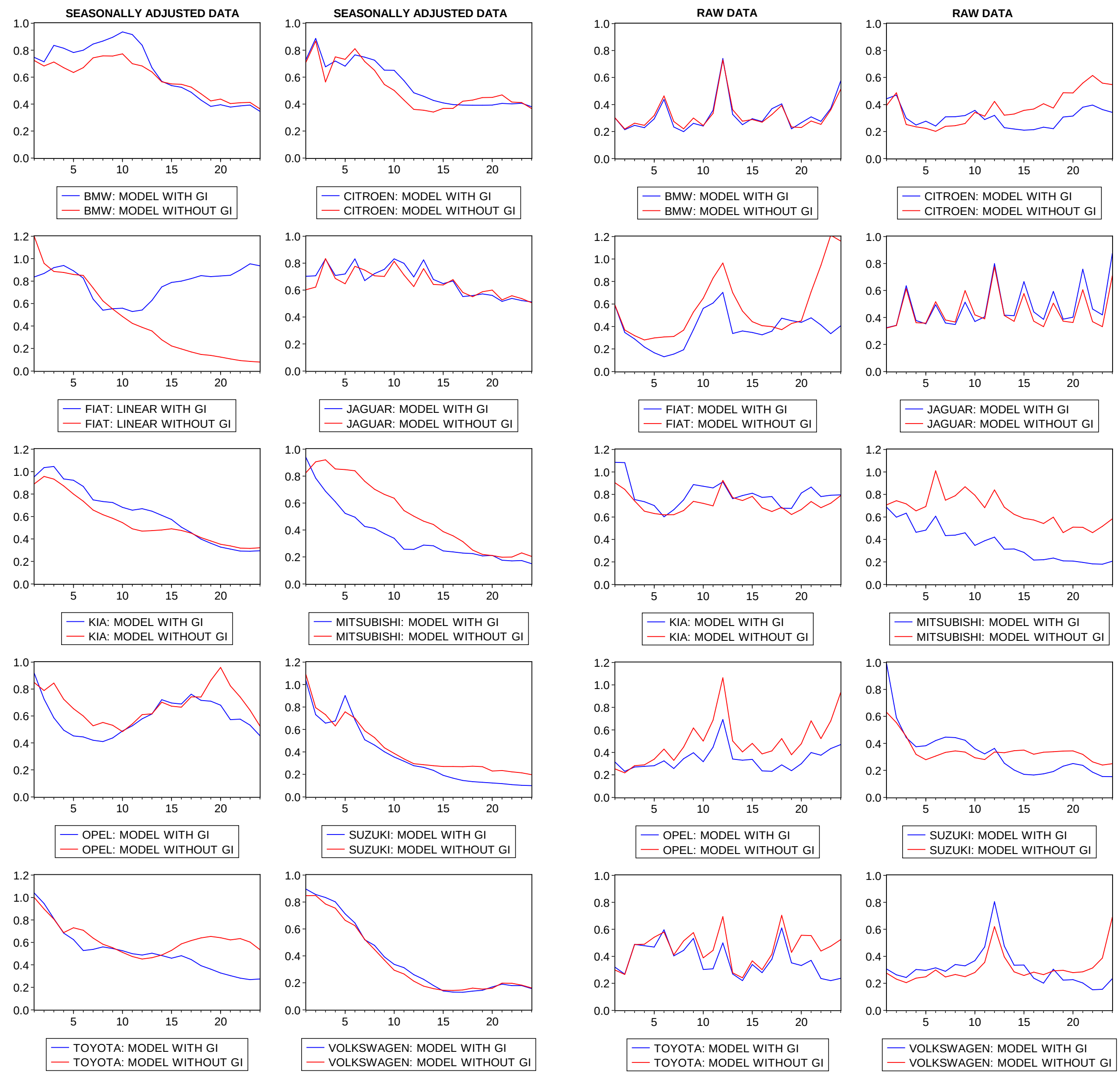

Figure 4: Ratios of the MSPEs of the best models with and without Google data and the Random Walk model across all forecasting horizons. The first two columns show results for seasonally-adjusted data, and the last two for raw data.

The ratios in Figure 4 show that it is difficult to outperform the random walk model in the case of short-term forecasts. Moreover, the best models without Google data tend to perform better than the best models with Google data for short and medium forecasts, whereas in general models using Google data show lower MSPEs for long-term forecasts with horizons higher than 12 steps ahead. This evidence suggests that potential gains in terms of forecasting performance may be achieved by using forecast combination methods. The development of these methods is beyond the scope of this paper and will be the subject of future studies.

Model rankings in terms of the MSPE do not show whether the competing forecasts are statistically different or not. We therefore tested for significant differences in forecast accuracy using the Model Confidence Set (MCS) approach proposed by Hansen, Lunde, and Nason (2011). The MCS is a sequential test of equal predictive ability, with the starting hypothesis that all models considered have equal forecasting performance. Given an initial set of forecasts, it tests the null that no forecast is distinguishable from any other and discards any inferior forecasts if they exist. The MCS procedure yields a model confidence set containing the best forecasting models at a given confidence level. Since our dataset is not too large and the number of forecasting models is moderate, we employed the semiquadratic test statistic (T_{SQ}), which is more computationally intensive but more selective, see e.g. Rossi and Fantazzini (2014). The loss function used was the MSPE, while the p-values for the test statistic were obtained using a stationary block bootstrap with a block length of 12 months and 1000 re-samples. If the p-value was lower than a defined confidence level α, the model was not included in the MCS and viceversa. A brief description of the MCS approach is reported in the Technical Appendix F.

The models included in the MCS at the 10% level for all car brands and forecast horizons are reported in Table 16[4]: for the sake of space and interest, we report only the total number of selected models, the total number of selected Google-based models, and whether the Random Walk model was included or not. The full set of results is available from the authors upon request.

Table 16 shows that most, if not all, models are selected in the case of forecasts up to 10-12 steps ahead for five car brands out of ten: the differences in forecasting performances are not large enough to distinguish between them, meaning that the MCS contains a large number of models. Moreover, the Random Walk model is often included. Instead, for long-term forecasts (12 steps ahead and higher), only a small number of models is selected, most of them bivariate models including only car sales and GIs, Bayesian VARs with GIs and sometimes the AR(12). Besides, the Random Walk model is seldom included. Here, the data are much more informative and it is possible to select a limited number of models which statistically outperform their competitors.

3.2 Raw data

We compared the same 34 models used for seasonally-adjusted data, but augmented with centered seasonal dummies to model potential seasonal behavior. Moreover, we also considered the bivariate Periodic Error Correction Model PECM(1,12) which includes only car sales and Google data, as discussed in section 2.2.2. To account for the possible endogeneity of regressors and improve the efficiency of the parameter estimates in small samples, we estimated the error correction term using the method of dynamic OLS (see Boswijk and Franses (1995), Hayashi (2000) and Franses and Paap (2004)). A short-cut notation for identifying each model in the subsequent tables reporting their forecasting performances is reported in the last column of Table 7.

We used the data between 2001M1 and 2009M6 as the first initialization sample for the models without Google data, while we used the initialization sample 2004M1-2009M6 for the models with Google data and for those without Google data but estimated on a shorter sample. The evaluation period ranged from 2009M7 till 2014M6 and was used to compare forecasts from 1 step ahead up to 24 steps ahead. The top three models in terms of the Mean Squared Prediction Error (MSPE) for each forecasting horizon and each car brand are reported in Table 17, while a summary of the models included in the MCS is reported in Table 18. The ratios of the MSPE of the best model with Google data and the Random Walk model across all forecasting horizons, together with the ratios of the MSPE of the best model without Google data and the Random Walk model are shown in the last two columns of Figure 4.

The results are somewhat similar to those which emerged from seasonally-adjusted data, but there are also some important differences. Models without Google data now perform better, with respect to the case of seasonally-adjusted data. Moreover, the number of models selected in the MCS is now much smaller (often no more than 2-6 models): Bayesian VARs (with and without Google data) and parsimonious bivariate models including only sales and GIs again represent the majority of models included in the MCS at the 10% level.

[4] We set $\alpha = 0.10$ as in Hansen, Lunde, and Nason (2011).

4 Robustness Checks

We wanted to verify that the superior performance of Google-based models also holds under alternative forecasting. We performed a series of robustness checks, considering alternative nonlinear models, alternative out-of-sample intervals, evaluating the directional accuracy of the competing forecasting models, checking whether Google data downloaded on different days can affect the models' forecasting performances, and examining additional car brands.

4.1 Nonlinear Models

A part of the economic and financial literature has suggested the use of nonlinear models for forecasting purposes (for instance, see Franses and Dijk (2000) and Terasvirta, Tjostheim, and Granger (2011) for a discussion at the textbook level). Given this evidence, we estimated a set of nonlinear models and compared their forecasting performances with the models in section 3. More specifically, we considered three nonlinear models:

- the SETAR model with 2 regimes (see Tong (1990) for a discussion at the textbook level);
- the logistic smooth transition autoregressive (LSTAR) model, which is a generalization of the SETAR model (see Tong (1990));
- the additive autoregressive model (AAR), also known as generalized additive model (GAM), since it combines generalized linear models and additive models (see Wood (2006) for a discussion at the textbook level).

A description of these nonlinear models is given in the Technical Appendix G. See D'Amuri and Marcucci (2013) and Fantazzini and Fomichev (2014) for a discussion of robustness checks using these nonlinear models.

The top three models in terms of the MSPE for each forecasting horizon and each car brand are reported in Table 19 for seasonally-adjusted data and in Table 21 for raw data. A summary of the models included in the MCS is reported in Table 20 for seasonally-adjusted data and in Table 22 for raw data.

In general, nonlinear models are very competitive, thus confirming past literature dealing with car sales forecasting (see Da, Engelberg, and Pengjie (2003), Kunhui, Qiang, Changle, and Junfeng (2007), Brühl, Borscheid, Friedrich, and Reith (2009), Hulsmann, Borscheid, Friedrich, and Reith (2012)). Particularly, parsimonious AAR and SETAR models involving only a few lags are often ranked among the top models in terms of MSPE. Moreover, AAR models with log-prices performed very well for medium- and long-term forecasts, similarly to what was found in Fantazzini and Fomichev (2014) when forecasting the real price of oil. However, nonlinear models were difficult to estimate, and specifications with a large number of lags failed to converge. Particularly, the LSTAR proved to be the most challenging and computationally intensive (see Franses and Dijk (2000) for a discussion of this issue). The results of the MCS confirm this evidence and most of the models included at the 10% level are nonlinear, whereas the only selected linear models are mostly Google-based. This evidence therefore seems to suggest that Google data may explain a good portion of the nonlinearity displayed by sales data.

In the case of raw data, nonlinear models are less competitive than linear models, particularly for forecasting horizons up to 12 steps ahead, whereas Bayesian VAR models and bivariate linear models including car sales and GIs are often the top ranked models across most of the car brands. However, for long-term forecasts, more than half of the models included in the MCS are nonlinear, while the remaining selected models are mainly bivariate Google-based models.

Tables 8-11 report the MSPEs, rankings, and eventual inclusion in the MCS of the best models in the case of 6, 12, 18, 24 step-ahead forecasts, respectively, for four model classes: linear models with GI, linear models without GI, nonlinear models and Random Walk models. Parsimonious bivariate models including only car sales and GIs are the best in the first class; AR(12) models and Bayesian models usually top the second class, while AAR and SETAR models with few lags are the best nonlinear models. The Random Walk has low rankings in long-term forecasts, but fares better for short-term forecasts.

SEASONALLY ADJUSTED DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	BVAR	7255188	36	Yes	VARD	1937867	4	Yes	VARongo112	11439630	10	Yes	VADongo112	7409	22	Yes	VARongo112	1220217	18	Yes
Linear w/o GI	AR12	6080991	4	Yes	VARDNOGO	2054348	6	No	BVARNOGO	11763929	11	Yes	AR12	6913	11	Yes	AR12	1036132	12	Yes
Nonlinear	AAR(7)dlog	5698093	1	Yes	LSTAR(3)log	1688104	1	Yes	SETAR(2)log	9458281	1	Yes	SETAR(5)log	6478	1	Yes	AAR(4)log	909397	1	Yes
RW		9073312	54	Yes		2534232	36	No		13839395	26	Yes		8907	61	Yes		1406331	28	Yes
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VECongo112	285910	1	Yes	VECongo112	16265040	1	Yes	VARongo112	1206941	9	Yes	VECongo112	4040621	1	Yes	VARongo112	60790600	2	Yes
Linear w/o GI	VECMNOGO4	484285	32	Yes	VECMXNOGO	21934786	6	No	VECMXNOGO	1233267	10	Yes	VECMXNOGO	5429724	3	Yes	AR124	58970200	1	Yes
Nonlinear	SETAR(11)log	347695	3	Yes	AAR(2)log	24373967	9	No	SETAR(3)log	812440	1	Yes	AAR(1)log	5653546	6	Yes	AAR(1)log	71730620	9	No
RW		577041	62	Yes		36590894	40	No		1760865	27	No		7663277	29	No		94350150	36	No
RAW DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VADongo112	9546849	1	Yes	VARongo112	913022	3	Yes	VARongo112	1225557	1	Yes	BVARPD	7852	1	Yes	BVARPD	731755	1	Yes
Linear w/o GI	BVARPDNOGO	10099266	2	Yes	BVARPNOGO	763696	1	Yes	BVARPNOGO	2874716	13	Yes	BVARPDNOGO4	8209	2	Yes	BVARPDNOGO	753717	3	Yes
Nonlinear	AAR(8)log	17435149	24	No	AAR(3)log	1121818	6	No	LSTAR(2)log	2615865	2	Yes	AAR(3)log	12342	30	No	AAR(9)log	828305	6	Yes
RW		21780883	62	No		3800014	70	No		9377139	62	No		15886	90	No		1218423	33	No
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VEPongo112	193905	1	Yes	VEPongo112	11342673	1	Yes	BVARP	523703	12	Yes	VEPongo112	3433042	2	Yes	VEPongo112	48248128	2	Yes
Linear w/o GI	BVARPDNOGO	322456	26	Yes	BVARPNOGO	15033303	5	No	BVARPNOGO4	380224	10	Yes	VECMXPNOGO4	3333000	1	Yes	BVARPNOGO	45734023	1	Yes
Nonlinear	LSTAR(3)log	234530	2	Yes	SETAR(9)dlog	31266013	29	No	SETAR(2)log	269229	1	Yes	SETAR(1)dlog	4165033	6	Yes	SETAR(4)log	72437319	12	No
RW		319260	24	Yes		34959425	38	No		1244803	34	No		5758188	30	No		153257703	63	No

Table 8: List of the best models for each model class, the corresponding MSPE and ranking. Forecast horizon: 6 steps ahead.

SEASONALLY ADJUSTED DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VARongo112	11630497	20	Yes	VECongo112	1746615	9	Yes	VARongo112	8070599	15	No	VECongo112	9089	15	Yes	VARongo112	1226001	24	
Linear w/o GI	AR12	9491265	5	Yes	BVARNOGO	1302593	6	Yes	BVARNOGO	5773194	10	No	AR12	8158	3	Yes	AR12	859701	14	
Nonlinear	SETAR(1)log	8506115	1	Yes	SETAR(4)log	1170021	1	Yes	AAR(1)log	2142114	1	Yes	SETAR(3)log	7758	1	Yes	SETAR(3)log	709664	1	
RW		13909104	43	Yes		3606090	43	No		14889930	30	No		13024	63	Yes		1830155	29	
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VECongo12	227696	1	Yes	VECongo112	21470315	3	Yes	BVAR	921548	12	No	BVAR	5198460	6	Yes	VARongo112	34715966	2	Yes
Linear w/o GI	AR12	446751	5	Yes	BVARNOGO	22609378	5	Yes	BVARNOGO4	979593	14	No	BVARNOGO4	4832954	1	Yes	AR124	28352989	1	Yes
Nonlinear	SETAR(11)log	393336	12	Yes	SETAR(7)dlog	20607165	1	Yes	AAR(3)log	222809	1	Yes	SETAR(7)dlog	4977505	2	Yes	SETAR(7)log	37548975	4	Yes
RW		888768	70	Yes		37088135	43	No		3319004	44	No		10675830	36	No		133013321	47	No
RAW DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VADongo112	11298025	3	Yes	VARongo112	872250	1	Yes	PECM	2688297	15	Yes	BVARPD	9021	2	Yes	VADongo112	859092	3	Yes
Linear w/o GI	AR124	11133748	1	Yes	AAR(3)log	1109012	3	Yes	BVARPNOGO4	3686731	23	No	BVARPDNOGO4	8749	1	Yes	BVARPDNOGO	871338	5	Yes
Nonlinear	SETAR(10)log	13316723	9	No	BVARPNOGO	1156313	4	Yes	SETAR(1)log	1732960	1	Yes	SETAR(7)dlog	11973	13	Yes	AAR(9)log	765639	1	Yes
RW		15260733	16	No		2734442	47	No		3822802	24	Yes		11294	9	Yes		942287	10	Yes
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VEPongo112	207765	1	Yes	VEPongo112	12940989	1	Yes	BVARP	585175	13	No	VEPongo112	2398770	1	Yes	BVARP	78930479	7	Yes
Linear w/o GI	AR12	414701	26	Yes	BVARPNOGO	19896671	4	No	BVARPNOGO4	541203	12	Yes	BVARPDNOGO4	3334235	6	Yes	BVARPNOGO	60857150	1	Yes
Nonlinear	AAR(1)log	273074	2	Yes	AAR(1)dlog	24701172	13	No	AAR(4)log	240509	1	Yes	SETAR(6)log	3292246	5	Yes	SETAR(6)log	72922923	2	Yes
RW		493513	41	Yes		18710646	3	Yes		1613001	38	No		4803964	26	No		98132979	22	Yes

Table 9: List of the best models for each model class, the corresponding MSPE and ranking. Forecast horizon: 12 steps ahead.

SEASONALLY ADJUSTED DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VARongo112	8464643	2	Yes	VECongo112	1920350	7	Yes	VECongo112	13037790	21	No	VECongo112	9462	2	Yes	VARongo112	940676	19	Yes
Linear w/o GI	AR12	9384237	3	Yes	BVARNOGO	2109563	9	Yes	AR124	2271788	4	Yes	AR12	9313	1	Yes	AR124	970566	21	Yes
Nonlinear	AAR(1)log	8018508	1	Yes	AAR(1)log	1266893	1	Yes	AAR(1)log	783045	1	Yes	LSTAR(5)log	10319	6	Yes	SETAR(8)log	780965	1	Yes
RW		19689847	42	No		4907445	38	No		15368830	27	No		16920	50	No		2358597	29	No
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VECongo112	174546	1	Yes	VECongo112	18394010	11	Yes	VARongo112	634907	9	No	VECongo112	3866818	1	Yes	VARongo112	14865460	1	Yes
Linear w/o GI	VECMNOGO4	195035	4	Yes	AR12	19024430	12	Yes	VECMXNOGO	1279649	15	No	BVARNOGO4	6302790	19	Yes	AR124	17283870	2	Yes
Nonlinear	SETAR(10)log	288480	5	No	SETAR(7)dlog	11826920	1	Yes	SETAR(2)log	191776	1	Yes	SETAR(3)dlog	4001607	2	Yes	AAR(1)log	26144200	4	Yes
RW		777574	66	No		25704740	29	Yes		4680547	37	No		9851512	36	No		107189500	43	No
RAW DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VARPD	15598460	11	Yes	VARongo112	1402351	2	Yes	VEPongo112	5156512	20	Yes	VEPongo112	13929	2	Yes	VEPongo112	1249080	5	Yes
Linear w/o GI	AR124	15171520	8	Yes	BVARPNOGO	2374012	15	Yes	AR12	4056693	18	Yes	VARPNOGO	11879	1	Yes	AR124	1262323	6	Yes
Nonlinear	LSTAR(9)log	12761820	1	Yes	AAR(3)log	1151259	1	Yes	LSTAR(2)log	2217764	1	Yes	AAR(3)log	14774	4	Yes	AAR(9)log	967518	1	Yes
RW		38548350	62	No		6342691	66	No		10898232	44	No		23496	67	Yes		1841665	29	No
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VEPongo112	159969	1	Yes	VARongo112	10159473	1	Yes	VARongo112	463961	13	Yes	VEPongo112	3849544	1	Yes	VECPongo12	58221750	2	Yes
Linear w/o GI	BVARPNOGO4	395301	24	Yes	BVARPDNOGO	18345564	3	Yes	BVARPNOGO4	816324	17	No	BVARPDNOGO	4438429	6	Yes	BVARPNOGO	55863450	1	Yes
Nonlinear	AAR(1)log	257037	4	Yes	SETAR(10)dlog	29080284	11	No	AAR(2)log	237811	1	Yes	SETAR(8)log	4084996	3	Yes	SETAR(2)log	83379030	10	Yes
RW		685032	60	No		35161323	16	No		2418065	36	No		6310415	29	No		190773000	38	No

Table 10: List of the best models for each model class, the corresponding MSPE and ranking. Forecast horizon: 18 steps ahead.

SEASONALLY ADJUSTED DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VARongo112	8099346	2	Yes	VARongo112	2770683	9	No	VECongo112	16937190	22	No	VECongo112	10310	2	Yes	VARongo112	965237	7	No
Linear w/o GI	AR124	8529251	3	Yes	AR124	2670345	8	No	AR124	1425390	2	Yes	AR12	10164	1	Yes	AR124	1056006	13	No
Nonlinear	AAR(1)log	7144053	1	Yes	AAR(1)log	1476196	1	Yes	AAR(1)log	906669	1	Yes	AAR(2)log	11325	5	Yes	SETAR(6)log	816700	1	Yes
RW		23423336	41	No		7290366	31	No		18093150	24	No		20103	41	No		3274180	30	No
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VECongo112	145162	1	Yes	VECongo112	13706410	2	Yes	VARongo112	711434	9	No	VECongo112	4005582	1	Yes	VARongo112	17513320	1	Yes
Linear w/o GI	VECMNOGO4	198324	3	Yes	VECMXNOGO	15895950	6	Yes	VECMXNOGO	1414340	13	No	VECMXNOGO	7803276	18	No	AR124	18042460	2	Yes
Nonlinear	SETAR(10)log	175800	2	Yes	SETAR(8)dlog	11794440	1	Yes	SETAR(2)log	208338	1	Yes	SETAR(6)dlog	5568402	2	Yes	AAR(1)log	27725190	4	Yes
RW		972338	62	No		30265060	32	No		7152681	34	No		14568460	42	No		111008400	40	No
RAW DATA																				
	Bmw				**Citroen**				**Fiat**				**Jaguar**				**Kia**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VADongo112	19015100	29	Yes	VARongo112	2029587	10	Yes	VARongo112	1822430	5	Yes	VEPongo112	15424	3	Yes	VARongo12	1256239	5	No
Linear w/o GI	AR12	16978200	10	Yes	BVARPNOGO	3248731	28	Yes	VARPDNOGO	5203134	24	No	VARPNOGO	12523	1	Yes	AR124	1244649	4	No
Nonlinear	LSTAR(8)log	12820190	1	Yes	AAR(3)log	987284	1	Yes	LSTAR(5)log	1600311	1	Yes	LSTAR(10)log	16592	11	Yes	AAR(9)log	1000641	1	Yes
RW		33076160	51	Yes		5935189	55	No		4488673	21	No		17635	17	Yes		1577128	16	No
	Mitsubishi				**Opel**				**Suzuki**				**Toyota**				**Volkswagen**			
	Model	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*	*Model*	*MSE*	*Ranking*	*MCS*
Linear w. GI	VARongo112	154734	1	Yes	VARongo112	9654207	1	Yes	VARongo112	640328	14	No	VARongo112	2192391	1	Yes	VECPongo12	23921080	1	Yes
Linear w/o GI	AR12	436153	42	No	VECMXPNOGO4	19153863	3	Yes	BVARPNOGO4	1044408	15	No	VECMPNOGO	4851103	7	Yes	BVARPNOGO	70306330	4	No
Nonlinear	AAR(1)log	216117	3	Yes	LSTAR(10)dlog	23307805	10	No	AAR(2)log	211698	1	Yes	SETAR(6)log	4730786	4	Yes	SETAR(3)log	75637050	5	No
RW		748481	65	No		20556112	7	Yes		4180500	47	No		9257854	37	Yes		101935900	21	No

Table 11: List of the best models for each model class, the corresponding MSPE and ranking. Forecast horizon: 24 steps ahead.

The previous evidence is confirmed and summarized by Figure 5 in Appendix C, which shows the ratios of the MSPEs of the best models with and without Google data with those of the Random Walk model, together with the ratios of the MSPEs of the best nonlinear models and the Random Walk model across all forecasting horizons: nonlinear models tend to perform better with seasonally adjusted data and medium- and long-term forecasts.

Finally, for the sake of interest (given the importance of long-term forecasts for car manufacturers) and space, we report in Tables 23 and 24 the list of models included in the Model Confidence Set for each car brand for 24 step-ahead forecasts, for seasonally-adjusted data and raw data, respectively. In the latter case, the number of models selected is higher on average than for seasonally-adjusted data, which was expected given the more noisy nature of raw data.

4.2 Alternative Out-of-Sample Periods

Our baseline out-of-sample interval includes the global financial crisis which started in 2007 and had a strong effect on car sales. Moreover, our in-sample analysis highlighted a potential structural break in the years 2008-2009. Therefore, we want to verify that our results continue to hold with different business cycle conditions, as recently highlighted by D'Amuri and Marcucci (2013). We considered the following two alternative out-of-samples:

- 2008M10-2009M6: this sample includes the official period of recession in Germany.
- 2009M7-2014M6: this sample starts after the end of the recession.

Due to the dimensionality of these new out-of-samples, we considered forecasts up to only 8 steps ahead. Moreover, this robustness check was performed only with seasonally-adjusted data, since the first forecast with raw data takes place after the end of the recession[5]. The top three models in terms of the MSPE for each forecasting horizon and each car brand are reported in Table 25 for the recession period, and in Table 26 for the expansion period.

The results are somewhat mixed and change substantially according to the car brand which is examined. However, some general indications can still be gained: Google-based models and linear models without Google data were the best models during the recession, while Google-based models and nonlinear models performed (slightly) better during the economic expansion. These results therefore provide further evidence of a structural break in the years 2008-2009. In general, Google-based models had forecasting performances which were more robust across different business cycles than their competitors, thus confirming similar evidence found by D'Amuri and Marcucci (2013) and Fantazzini and Fomichev (2014).

4.3 Directional Accuracy

The analysis has so far only considered the accuracy of forecasts in terms of magnitude, but directional accuracy is also important: forecasts with the correct direction of change may still provide useful information even with large forecast errors. This is particularly important when predicting a turning point, which is a special case of directional accuracy and represents a change in the direction of movement of the analyzed variable (Theil (1961) and Naik and Leuthold (1986)).

The top three models in terms of average directional accuracy (in %) for each car brand, for short-term forecasts (1-6 steps ahead), medium-term forecasts (7-12 steps ahead), and long-term forecasts (13-24 steps ahead) are reported in Table 27 (top part) for seasonally adjusted data and in Table 27 (bottom part) for raw data.

In the case of seasonally-adjusted data, parsimonious bivariate models, including only car sales and GIs, as well as AAR models had the higher percentage of correct forecasts of the direction of change for most of the car brands and forecasting horizons. As for raw data, similarly to what we saw in section 4.1, nonlinear models are, in general, less competitive than linear models. More specifically, linear models without Google data performed better than with seasonally-adjusted data (particularly for short-term directional forecasts), while nonlinear models were competitive only for medium- to long-term directional accuracy. Instead, Google-based models performed relatively well and simple bivariate models with car

[5]Raw data required a larger initialization sample due to the inclusion of centered seasonal dummies.

sales, GIs and centered seasonal dummies provided very precise forecasts of the direction of change for most of the car brands and forecasting horizons.

The somewhat differing results between seasonally-adjusted data and raw data could be due to two reasons. Firstly, the procedure of seasonal adjustment changes the statistical properties of the data and can affect considerably the models' forecasting performances (Zellner (1978) and Franses and Paap (2004)). Secondly, Boivin and Ng (2006) and Stock and Watson (2006) have shown that small models may outperform models with a larger number of parameters because they allow for a better extraction of relevant signals than models overloaded with parameters and complex specifications. In this regard, Google data allow us to summarize a lot of information and reduce model complexity.

4.4 Sampling Variability of Google Data

Google data does not refer to the population of searches, but only to a sample. As a consequence, the time series of Google data can vary substantially from one download to another[6]. We downloaded the GIs for a number of subsequent days to check how sampling variability can affect the models' forecasting performances. More specifically, we compared the forecasts computed in our baseline case with GIs downloaded on 15/08/2014, with forecasts computed with the average GIs downloaded between the 15/08/2014 and the 02/09/2014. We used the average GIs following the approach recently proposed by Carriere-Swallow and Labbé (2013). Table 28 shows the average ratio – averaged across all forecasting horizons – of the MSPE for the forecasts computed with GIs downloaded on the 15/08/2014, with respect to MSPE for the forecasts computed with the average GIs downloaded between 15/08/2014 and 02/09/2014.

Almost all models have ratios close to 1, with the notable exception of high-dimensional VEC models, which did not reach convergence for a couple of car brands (Toyota and Kia). The large variance of estimators for cointegrated models in small-medium samples is a well known issue in the econometric literature (Stock and Watson (1993), Maddala and Kim (1998)(section 5.7) and Hayashi (2000)(section 10.4)): most likely, the sampling noise of Google data exacerbates this inference problem. Using average GIs can solve this issue to some extent, but not completely: the high-dimensional VECM models still did not reach convergence in some cases. Moreover, the rankings of Google-based models in the case of averaged data are very close, if not identical, to the rankings of Google-based models in the baseline case for all car brands (results not reported). Therefore, the most advisable solution is probably either to use parsimonious VEC models or revert to Bayesian methods.

4.5 Additional Car Brands

In the baseline section, we analyzed 10 car brands out of the 22 car brands which both have monthly data continuously available since 2001 and are present in Google Trends. We briefly examine here the forecasting performances of the remaining 12 car brands:

- Large sellers: Ford, Audi;
- Medium-sized sellers: Hyundai, Mazda, Nissan, Peugeot, Renault;
- Small sellers: Honda, Land Rover, Porsche, Subaru, Volvo.

Table 29 and 30 report the top three models in terms of MSPE for each forecasting horizon and each car brand, in the case of seasonally-adjusted data and raw data, respectively.

The results are similar to those of the baseline case: parsimonious bivariate linear models involving only GIs and car sales and nonlinear models (with few lags) are the best models for all brands examined. Bayesian models are a valid alternative for short-term forecasting in the case of seasonally-adjusted data.

5 Conclusions

This paper proposed a set of multivariate models for forecasting car sales using both Google data and economic variables. Moreover, we considered multivariate models for both deseasonalized data and for raw data. We performed a forecasting exercise for ten car brands in Germany, and we computed

[6]The authors want to thank an anonymous referee for pointing out this issue.

out-of-sample forecasts ranging from 1 month to 24 months ahead. Our results showed that Bayesian VAR models performed rather well for all car brands and for short- and medium-term forecasts, while parsimonious bivariate models including only car sales and Google models outperformed the competing models in the case of long-term forecasts for several brands. Furthermore, the forecasting power of the best Google-based models increased with the length of the forecast horizon, particularly with forecast horizons higher than 12 steps ahead. Apart from this, no particular differences between large, medium-sized and small sellers and between foreign and German manufacturers were found. In case of raw data, models without Google data performed better than in the case of seasonally-adjusted data. However, Bayesian VARs (with and without Google data) and parsimonious bivariate models including only sales and Google data represented again the majority of models included in the MCS at the 10% level. Finally, we performed a set of robustness checks to verify that our results also hold under different forecasting setups. We found out that nonlinear AAR and SETAR models were very competitive and were included in the MCS together with Google-based models, thus suggesting that Google data may explain a part of the nonlinearity displayed by sales data. However, nonlinear models were difficult to estimate and on several occasions failed to converge. Alternative out-of-sample intervals highlighted that Google-based models performed better during the recession (which is of particular importance for car manufacturers) and, in general, they had forecasting performances which were more robust across different business cycles than their competitors. Our previous results also held in the case of directional accuracy, which showed that Google-based models provided the most precise forecasts of the direction of change. We found that the sampling variability of Google data can be problematic for high-dimensional VEC models. Using the averaged Google data over several days can solve this issue to some extent, but parsimonious VEC models and Bayesian methods are valid alternatives as well. The results in the baseline case also held for twelve additional car brands.

Even though we considered a very large set of models, we had to restrict their potential range in order to keep the forecasting exercise computationally tractable. An avenue of future research would be to consider additional models such as fractional cointegration, exponential smoothing methods in state space form, and many others.

References

Aparicio, F., A. Escribano, and A. Garcia (2006): "Range Unit-Root (RUR) tests: robust against nonlinearities, error distributions, structural breaks and outliers," *Journal of Time Series Analysis*, 27(4), 545–576.

Askitas, N., and K. Zimmermann (2009): "Google econometrics and unemployment forecasting," *Applied Economics Quarterly*, 55, 107–120.

Boivin, J., and S. Ng (2006): "Are more data always better for factor analysis?," *Journal of Econometrics*, 132(1), 169–194.

Boswijk, H. (1994): "Testing for an unstable root in conditional and structural error correction models," *Journal of Econometrics*, 63, 37–60.

Boswijk, H., and P. Franses (1995): "Periodic cointegration–representation and inference," *Review of Economics and Statistics*, 77, 436–454.

——— (1996): "Unit roots in periodic autoregressions," *Journal of Time Series Analysis*, 17, 221–245.

Brühl, B.and Hülsmann, M., D. Borscheid, C. Friedrich, and D. Reith (2009): "A sales forecast model for the German automobile market based on time series analysis and data mining methods.," in *Advances in Data Mining Applications and Theoretical Aspects*, ed. by P. Perner, pp. 146–160. Springer.

Carriere-Swallow, Y., and F. Labbé (2013): "Nowcasting with Google Trends in an emerging market," *Journal of Forecasting*, 32(4), 289–298.

Choi, H., and H. Varian (2012): "Predicting the present with Google Trends," *Economic Record*, 88, 2–9.

Da, Z., J. Engelberg, and G. Pengjie (2003): "A comparative study of linear and nonlinear models for aggregate retail sales forecasting," *International Journal of Production Economics*, 86, 217–231.

——— (2011): "In Search of Attention," *Journal of Finance*, 5, 1461–1499.

D'Amuri, F., and J. Marcucci (2013): "The predictive power of Google searches in forecasting unemployment," Temi di discussione (Economic working papers) 891, Bank of Italy.

Dekimpe, M. G., D. M. Hanssens, and J. M. Silva-Risso (1998): "Long-run effects of price promotions in scanner markets," *Journal of Econometrics*, 89(1), 269–291.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews*, 3, 1–100.

Engle, R., and C. Granger (1987): "Co-integration and error correction: representation," *Econometrica*, 55, 251–276.

Fantazzini, D., and N. Fomichev (2014): "Forecasting the real price of oil using online search data," *International Journal of Computational Economics and Econometrics*, 4(1-2), 4–31.

Franses, P., and V. Dijk (2000): *Nonlinear time series models in empirical finance*. Cambridge University Press.

Franses, P., and R. Paap (1994): "Model selection in periodic autoregressions," *Oxford Bulletin of Economics and Statistics*, 56, 421–440.

——— (2004): *Periodic time series models*. Oxford University Press, Oxford.

Franses, P. H., and N. Haldrup (1994): "The effects of additive outliers on tests for unit roots and cointegration," *Journal of Business and Economic Statistics*, 12, 471–478.

Ginsberg, J., M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant (2009): "Detecting influenza epidemics using search engine query data," *Nature*, 457, 1012–1014.

Gregory, A., and B. Hansen (1996): "Residual-based tests for cointegration in models with regime shifts," *Journal of Econometrics*, 70, 99–126.

GTAI (2014): http://www.gtai.de.

Hall, S., S. Henry, and J. Greenslade (2002): "On the identification of cointegrated systems in small samples: A modelling strategy with an application to UK," *Journal of Economic Dynamics and Control*, 26, 1517–1537.

Hansen, P., A. Lunde, and J. Nason (2011): "The model confidence set," *Econometrica*, 79(2), 453–497.

Hartigan, J. A. (1975): *Clustering algorithms*. John Wiley & Sons, Inc., New York, NY, USA, 99th edn.

Hatemi, J. (2008): "Tests for cointegration with two unknown regime shifts with an application to financial market integration," *Empirical Economics*, 35, 497–505.

Hayashi, F. (2000): *Econometrics*. Princeton University Press, Princeton.

Hulsmann, M., D. Borscheid, C. M. Friedrich, and D. Reith (2012): "General sales forecast models for automobile markets and their analysis," *Transactions on Machine Learning and Data Mining*, 5(2), 65–86.

International Telecommunications Union (2014): "Internet users' graphs, market information and statistics," http://www.itu.int/ITU-D/ict/statistics.

Internet World Stats (2014): http://www.internetworldstats.com/stats.htm.

Johansen, S. (1995): *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, Oxford.

——— (2006): "Cointegration: a survey," in *Palgrave Handbook of Econometrics: Volume 1, Econometric Theory*, ed. by T. Mills, and K. Patterson, pp. 540–577. Palgrave MacMillan, Basingstoke, UK and New York, USA.

Johansen, S., R. Mosconi, and B. Nielsen (2000): "Cointegration analysis in the presence of structural breaks in the deterministic trend," *Econometrics Journal*, 3, 216–249.

Koop, G., and D. Korobilis (2010): "Bayesian multivariate time series methods for empirical macroeconomics," *Foundations and Trends in Econometrics*, 3(4), 267–358.

Kunhui, L., L. Qiang, Z. Changle, and Y. Junfeng (2007): "Time series prediction based on linear regression and SVR," in *Third International Conference on Natural Computation (ICNC2007)*, pp. 688–691, Haikou,China.

Landwehr, J. R., A. A. Labroo, and A. Herrmann (2011): "Gut liking for the ordinary: incorporating design fluency improves automobile sales forecasts," *Marketing Science*, 30, 416–429.

Lee, L., and M. Strazicich (2001): "Break point estimation and spurious rejections with endogenous unit root tests," *Oxford Bulletin of Economics and Statistics*, 63(5), 535–558.

——— (2003): "Minimum lagrange multiplier unit root test with two structural breaks," *Review of Economics and Statistics*, 85(4), 1082–1089.

Lin, Y., and W. J. Zhang (2004): "Aesthetic design for automobile interiors: critical and conceptual framework," in *2004 IEEE International Conference on Systems, Man and Cybernetics*, pp. 6313–6317, The Hague.

Litterman, R. (1986): "Forecasting with bayesian vector autoregressions: five years of experience," *Journal of Business and Economic Statistics*, 4, 25–38.

Lumsdaine, R., and D. Papell (1997): "Multiple Trend Breaks and the Unit-Root Hypothesis," *Review of Economics and Statistics*, 79(2), 212–218.

Lutkepohl, H. (2005): *New introduction to multiple time series analysis*. Springer, Berlin.

Maddala, G. S., and I. Kim (1998): *Unit Roots, cointegration, and structural change*. Cambridge University Press, York. ΩVDA

VDA (2014): http://www.vda.de.

Naik, G., and R. Leuthold (1986): "A note on qualitative forecast evaluation," *American Journal of Agricultural Economics*, 68, 721–726.

Net Applications (2014): "Market share statistics for internet technologies," www.Marketshare.hitslink.com.

Ng, S., and P. Perron (1995): "Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag," *Journal of the American Statistical Association*, 90, 268–281.

Nijs, V. R., M. G. Dekimpe, J.-B. E. M. Steenkamp, and D. M. Hanssens (2001): "The category-demand effects of price promotions," *Marketing Science*, 20, 1–22.

Nunes, L., P. Newbold, and C. Kuan (1997): "Testing for unit roots with breaks: evidence on the great crash and the unit root hypothesis reconsidered," *Oxford Bulletin of Economics and Statistics*, 59(4), 435–448.

Osterwald-Lenum, M. (1992): "A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics," *Oxford Bulletin of Economics and Statistics*, 54, 461–471.

Pauwels, K., D. M. Hanssens, and S. Siddarth (2002): "The long-term effects of price promotions on category incidence, brand choice, and purchase quantity," *Journal of Marketing Research*, 39, 421–439.

Pauwels, K., J. Silva-Risso, S. Srinivasan, and D. M. Hanssens (2004): "New products, sales promotions, and firm value: the case of the automobile industry," *The Journal of Marketing*, 68, 142–156.

Perron, P. (1990): "Testing for a unit-root in a time series with a changing mean," *Journal of Business and Economic Statistics*, 8, 153–162.

——— (2006): "Dealing with structural breaks," in *Palgrave Handbook of Econometrics - Vol.1:Econometric Theory*, ed. by K. Patterson, and T. Mills, pp. 278–352. Palgrave Macmillan.

Perron, P., and T. J. Volgelsang (1992): "Nonstationary time series and level shifts with an application to purchasing power parity," *Journal of Business and Economic Statistics*, 10, 301–320.

Pesaran, M., and Y. Shin (1998): "Impulse response analysis in linear multivariate models," *Economics Letters*, 58, 17–29.

Phillips, P., and S. Ouliaris (1990): "Asymptotic properties of residual based tests for cointegration," *Econometrica*, 58, 165–193.

Rossi, E., and D. Fantazzini (2014): "Long memory and Periodicity in Intraday Volatility," *Journal of Financial Econometrics*, forthcoming.

Sa-ngasoongsong, A., and S. T. S. Bukkapatnam (2011): "Willingness-to-pay prediction based on empirical mode decomposition," in *Proceedings of the 2011 Industrial Engineering Research Conference*, ed. by T. Doolen, and E. V. Aken, Reno.

Sa-ngasoongsong, A., S. T. S. Bukkapatnam, J. Kim, P. S. Iyer, and R. P. Suresh (2012): "Multi-step sales forecasting in automotive industry based on structural relationship identification," *International Journal of Production Economics*, 140, 875–887.

Shahabuddin, S. (2009): "Forecasting automobile sales," *Management Research News*, 32, 670–682.

Stock, J., and M. Watson (1993): "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica*, 61, 783–820.

Stock, J. H., and M. W. Watson (2006): "Forecasting with many predictors," *Handbook of economic forecasting*, 1, 515–554.

Suhoy, T. (2009): "Query indices and a 2008 downturn," Discussion Paper 6, Bank of Israel.

Terasvirta, T., D. Tjostheim, and C. Granger (2011): *Modelling nonlinear economic time series*. Oxford University Press, Amsterdam.

Theil, H. (1961): *Economic forecasts and policy*. North Holland Publishing Company, Amsterdam.

Tong, H. (1990): *Non-linear time series: a dynamical system approach*. Oxford University Press, Oxford.

Wang, F., K. Chang, and C. Tzeng (2011): "Using adaptive network-based fuzzy inference system to forecast automobile sales," *Expert Systems with Applications*, 38, 10587–10593.

Wood, S. (2006): *Generalized additive models: an introduction with R,*. Chapman and Hall / CRC, Boca Raton.

Zellner, A. (1978): *Seasonal analysis of economic time series*, no. zell78-1 in NBER Books. US Department of Commerce, Bureau of the Census.

Zivot, E., and K. Andrews (1992): "Further evidence on the great crash, the oil price shock, and the unit root hypothesis," *Journal of Business and Economic Statistics*, 10(10), 251–270.

A In-sample Analysis

	BC	CCI	CPI	EURIBOR	PP	GDP	PI	UR	Google
BMW	V	V	V		V	V	V	V	V
Citroen	V	V	V	V	V	V	V	V	V
Fiat	V	V	V	V	V	V	V	V	V
Jaguar	V	V	V		V	V	V	V	V
Kia	V	V	V		V	V	V	V	V
Mitsubishi	V		V				V	V	V
Opel	V	V	V	V	V	V	V	V	V
Suzuki	V	V	V	V	V	V	V	V	V
Toyota	V	V	V	V	V	V	V	V	V
Volkswagen	V	V	V	V	V	V	V	V	V

Table 12: Weak exogeneity of seasonally-adjusted data: variables for which the null hypothesis of weak exogeneity can be rejected after re-testing at the 5% probability level.

	BC	CCI	CPI	EURIBOR	PP	GDP	PI	UR	Google
BMW	V	V	V		V	V	V	V	V
Citroen	V	V	V	V	V	V	V	V	V
Fiat	V	V	V	V	V	V	V	V	V
Jaguar	V		V					V	V
Kia	V	V	V	V	V	V	V	V	V
Mitsubishi	V	V	V	V	V	V	V		V
Opel	V	V	V	V	V	V	V	V	V
Suzuki	V	V	V	V	V	V	V	V	V
Toyota	V	V	V	V	V	V	V	V	V
Volkswagen	V	V	V	V	V	V	V	V	V

Table 13: Weak exogeneity of raw data: variables for which the null hypothesis of weak exogeneity can be rejected after re-testing at the 5% probability level.

	Sample: 2001-2014		Sample: 2008-2014	
	Boskwijk joint Wald test statistic	**Weak exogeneity test of GIs (*p*-value)**	**Boskwijk joint Wald test statistic**	**Weak exogeneity test of GIs (*p*-value)**
BMW	46.90*	0.86	39.08*	0.85
Citroen	45.75*	0.63	58.38*	0.30
Fiat	65.87*	0.15	116.11*	**0.00**
Jaguar	22.26	0.05	19.83	0.49
Kia	48.45*	0.06	31.83*	0.95
Mitsubishi	37.46*	0.66	29.17	0.62
Opel	49.26*	0.53	29.70	0.54
Suzuki	41.72*	0.61	67.46*	0.25
Toyota	69.64*	**0.01**	47.02*	**0.02**
Volkswagen	41.10*	0.56	48.33*	0.37

Table 14: The null hypothesis is the absence of cointegration against the alternative of periodic cointegration. The Boskwijk test considered the case with seasonal intercepts. * Significance at the 5% level. p-values smaller than 5% are reported in bold.

B Forecasting performances: Baseline case

B.1 Seasonally Adjusted data

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12
BMW	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12
	BVARNOGO4	BVAR	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124
	BVAR	AR124	BVAR	BVAR	BVAR	BVAR	VADongo1112	VADongo1112	BVAR	BVARDNOGO4	VADongo1112	VARongo1112
OPEL	BVARDNOGO	VADongo12	VADongo12	VECongo1112	VECongo1112	VECongo1112	VECM	VECongo1112	VECongo1112	VECMXNOGO	VECongo1112	VECongo1112
	BVARDNOGO4	VECongo1112	VECongo1112	VADongo12	VECM	VECM	VECMX	VARongo1112	VARongo1112	VECMNOGO	VECMXNOGO	BVARNOGO
	BVARD	VARongo12	VARongo12	VECongo12	VECMX	VECMX	VECongo1112	VECMXNOGO	VECMXNOGO	VECongo1112	VECMNOGO	AR12
VOLKSWAGEN	BVARDNOGO	BVARDNOGO	VARNOGO	VARDNOGO	AR124	AR124	VARongo1112	AR124	AR124	AR124	AR124	AR124
	BVARDNOGO4	BVARD	VARXNOGO	VARXDNOGO	VARongo1112	VARongo1112	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARD	BVARDNOGO4	VARDNOGO	VARNOGO	VECongo1112	AR12	AR12	AR12	AR12	AR12	AR12	AR12
CITROEN	VARDNOGO4	VECMXNOGO4	VECMXNOGO4	VARD	VARD	VARD	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO
	VARXDNOGO4	VECMNOGO4	VECMNOGO4	VARXD	VARXD	VARXD	VARD	VARongo1112	VARongo1112	VARongo1112	VECongo1112	VECongo1112
	VECMXNOGO4	BVARDNOGO	VARD	VARDNOGO	VARDNOGO	VARDNOGO	VARXD	VAR	BVAR	VECongo1112	VARongo1112	VARongo1112
FIAT	VECongo1112	VADongo1112	AR12	AR12	AR12	VARongo1112	VARongo1112	VARongo1112	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO
	VARongo1112	VECongo1112	VADongo1112	VADongo1112	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO	VARongo1112	VARongo1112	AR12	AR124
	VADongo1112	AR12	VECongo1112	VECongo1112	VECongo1112	AR12	AR12	AR12	AR12	AR12	AR124	AR12
TOYOTA	RW	VECMXNOGO	VARNOGO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	BVAR	BVARNOGO4	BVARNOGO4	BVARNOGO4
	VARNOGO	VECMNOGO	VARXNOGO	VARNOGO	VARNOGO	BVAR	BVAR	BVAR	BVARNOGO4	BVAR	BVAR	BVAR
	VARXNOGO	VECongo1112	VECongo1112	VARXNOGO	VARXNOGO	VECMXNOGO	BVARNOGO4	BVARNOGO4	VECongo1112	VECongo1112	VECongo1112	VECongo1112
JAGUAR	AR12	AR12	VADongo1112	AR12	AR12	AR12	VECongo1112	AR12	AR12	AR12	AR12	AR12
	AR124	AR124	AR12	VECongo1112	AR124	VADongo1112	AR12	VECongo1112	VECongo1112	VECongo1112	AR124	AR124
	BVARDNOGO	BVARDNOGO	BVARDNOGO4	AR124	VECongo1112	BVARDNOGO4	AR124	AR124	BVARDNOGO4	BVARD	VECongo1112	VECongo1112
KIA	BVARDNOGO	AR12	AR124	BVARNOGO4	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12
	BVARDNOGO4	AR124	VARDNOGO4	AR12	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124
	BVARD	BVARDNOGO	AR12	AR124	BVARNOGO4	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
MITSUBISHI	BVARDNOGO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo12	VECongo12	VECongo12
	BVARDNOGO4	BVARNOGO4	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VECongo1112	VARongo1112	VECongo1112
	BVARNOGO4	VADongo1112	VADongo1112	VADongo1112	VECongo12	VECongo12	VARongo12	VARongo12	VARongo12	VARongo1112	VARongo12	VARongo1112
SUZUKI	RW	VARongo1112	VARongo1112	VECMXNOGO	VECMXNOGO	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VECongo1112	VECongo1112	BVAR
	RW4	BVARNOGO	BVARNOGO	VECMNOGO	VECMNOGO	VECMXNOGO	BVAR	VECongo1112	VECongo1112	BVARNOGO	BVAR	VECongo1112
	VARongo1112	BVARNOGO4	BVARNOGO4	VARongo1112	BVARNOGO	VECMNOGO	BVARNOGO	BVARNOGO	BVARNOGO	BVAR	BVARNOGO4	BVARNOGO4

	Step 13	Step 14	Step 15	Step 16	Step 17	Step 18	Step 19	Step 20	Step 21	Step 22	Step 23	Step 24
BMW	AR12	AR12	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	AR124	VARongo1112	AR12	AR12	AR12	AR12	AR124	AR124	AR124	AR124	AR124	AR124
	VARongo1112	AR124	AR124	AR124	AR124	AR124	AR12	AR12	AR12	AR12	AR12	AR12
OPEL	VECongo1112	AR12	AR12	AR12	AR12	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	AR12	VECongo1112	VECongo1112	VECongo1112	VECongo1112	AR12	VADongo12	VARongo1112	VECMXNOGO	VECMXNOGO	VECMXNOGO	VECMXNOGO
	BVARNOGO	AR124	AR124	AR124	AR124	VADongo1112	AR12	VADongo1112	VECMNOGO	VECMNOGO	VECMNOGO	VECMNOGO
VOLKSWAGEN	AR124	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	AR12	VARongo1112	AR124	AR124	AR124	AR124	AR124	AR12	AR124	AR124	AR124	AR124
	VARongo1112	AR12	AR12	AR12	AR12	AR12	AR12	VARongo1112	AR12	AR12	AR12	AR12
CITROEN	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	AR124
	VECongo1112	VECongo1112	VECongo1112	VECongo1112	BVARNOGO	BVARNOGO	BVARNOGO	BVARNOGO	VARongo1112	AR124	AR124	VARongo1112
	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AR124	VARongo1112	VARongo1112	VECongo1112
FIAT	BVARNOGO	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124	AR124
	AR124	BVARNOGO	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12
	AR12	AR12	BVARNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO
TOYOTA	BVARNOGO4	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	VECongo1112	BVARNOGO4	BVARNOGO4	BVARNOGO4	BVARNOGO4	BVARNOGO4	AR12	AR12	AR12	VECMXNOGO	VECMXNOGO	VECMXNOGO
	BVAR	BVAR	BVAR	BVAR	BVAR	AR124	BVARNOGO4	BVARNOGO4	AR124	VECMNOGO	VECMNOGO	VECMNOGO
JAGUAR	AR12	AR12	AR12	VECongo1112	VECongo1112	AR12	VECongo1112	VECongo1112	VARongo1112	VARongo12	VECongo1112	AR12
	AR124	VECongo1112	VECongo1112	AR12	AR12	VECongo1112	AR12	AR12	VARongo12	AR12	AR12	VECongo1112
	VECongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AR124	VARongo1112	AR12	VARongo1112	VARongo1112	AR124
KIA	AR12	AR12	AR12	AR124	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	AR124	AR124	AR124	AR12	VARongo1112	AR124	AR124	AR124	AR124	AR124	AR124	AR124
	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AR12	AR12	AR12	AR12	AR12	AR12	AR12	AR12
MITSUBISHI	VECongo1112	VECongo1112	VARongo1112	VARongo1112	VARongo1112	VECongo1112	VECongo1112	VECMNOGO4	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	VARongo1112	VARongo1112	VECongo1112	VECongo1112	VECongo1112	VARongo1112	VECMNOGO4	VARongo12	VECMNOGO4	VECMNOGO4	VECMNOGO4	VECMNOGO4
	VECongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo1112	VECongo1112	VARongo1112	VARongo1112	VARongo12	VARongo12
SUZUKI	BVAR	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARNOGO4	BVAR	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	BVARNOGO	BVARNOGO4	BVAR	BVAR	BVAR	BVAR	BVAR	VECMXNOGO	VECMXNOGO	VECMXNOGO	VECMXNOGO	VECMXNOGO

Table 15: Top three models in terms of MSPE for each forecasting horizon and each car brand.

BMW	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	15	34	32	34	34	33	22	21	22	32	31	31	28	21	23	22	6	7	6	6	6	6	6	7
Google models	6	14	13	14	14	13	10	10	10	13	13	13	13	10	12	11	3	3	3	3	3	3	3	3
Random Walk included?	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no
OPEL	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	5	4	7	8	2	2	1	5	5	1	2	4	8	9	9	8	9	10	9	8	4	2	2	2
Google models	1	3	6	6	2	2	1	3	3	0	1	1	4	3	4	3	4	4	4	4	3	1	1	1
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	no	no	no	no
VOLKSWAGEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	6	4	1	1	3	4	2	1	2	2	2	2	2	3	5	5	4	3	4	4	4	3	5	4
Google models	2	1	0	0	1	1	1	0	1	1	1	1	1	1	3	3	2	1	2	2	2	1	3	2
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
CITROEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	1	1	1	34	34	34	2	4	4	4	4	4	4	4	5	7	10	9	7	8	9	9	13	13
Google models	0	0	0	14	14	14	0	1	2	2	2	2	2	2	3	4	4	4	3	4	5	5	6	6
Random Walk included?	no	no	no	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no
FIAT	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	6	7	8	11	12	16	10	10	11	9	6	5	4	4	4	4	4	2	2	1	1	1	1	1
Google models	3	2	3	4	4	7	4	3	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0
Random Walk included?	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no
TOYOTA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	2	1	1	2	2	3	4	4	4	4	5	5	6	7	7	11	9	2	2	2	1	2	2	3
Google models	0	0	0	1	1	2	2	2	2	2	2	2	2	3	3	5	3	1	1	2	1	1	1	2
Random Walk included?	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no
JAGUAR	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	27	34	9	21	22	24	22	23	29	22	11	15	17	7	5	6	4	7	4	1	7	6	7	7
Google models	11	14	4	9	11	10	9	10	14	10	5	8	9	4	3	3	3	5	2	1	4	3	4	4
Random Walk included?	no	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no
KIA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	20	32	34	31	26	28	33	34	34	23	25	10	21	19	19	19	16	10	10	8	8	34	30	4
Google models	8	13	14	14	10	11	14	14	14	10	10	5	12	9	9	9	8	6	5	3	3	14	13	1
Random Walk included?	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no
MITSUBISHI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	28	33	34	33	33	34	32	31	7	11	9	8	11	11	6	4	4	5	5	7	6	5	8	9
Google models	10	14	14	14	14	14	14	13	5	6	6	6	7	7	4	3	3	4	3	4	4	3	5	6
Random Walk included?	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
SUZUKI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	9	9	3	1	1	2	6	6	6	6	7	7	7	6	5	7	5	4	2	2	2	2	2	2
Google models	4	4	1	0	0	1	3	3	3	3	4	4	4	3	3	4	3	3	2	2	2	2	2	2
Random Walk included?	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no

Table 16: Models included in the Model Confidence Set for each forecast horizon (from 1 step up to 24 steps ahead) and for each car brand.

B.2 Raw data

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12
BMW	BVARPNOGO4	BVARPD	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	AR12	AR124
	BVARP	BVARP	BVARPDNOGO	BVARPDNOGO	AR12	BVARPDNOGO	BVARPD	BVARPD	BVARPDNOGO	AR12	AR124	AR12
	BVARPNOGO	BVARPDNOGO	BVARPDNOGO4	BVARPD	BVARPDNOGO	AR12	BVARPDNOGO4	BVARPDNOGO4	AR12	AR124	VADongo1112	VADongo1112
OPEL	BVARPDNOGO	BVARPNOGO	VARongo1112	VARP	VARP	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPNOGO	VEPongo1112	VEPongo1112	VARXP	VARXP	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VADongo1112	VADongo1112
	BVARPDNOGO4	BVARPDNOGO	BVARPD	VEPongo1112	VEPongo1112	VARP	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VARongo1112	RW
VOLKSWAGEN	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO
	VECMXPNOGO	BVARP	VARPNOGO	VARPNOGO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	BVARP	VEPongo1112	VEPongo1112	BVARP
	VECMPNOGO	BVARPDNOGO	VARXPNOGO	VARXPNOGO	VARPNOGO	VADongo1112	VARPNOGO	BVARP	VEPongo1112	BVARP	VADongo1112	VADongo1112
CITROEN	VARPNOGO4	VARongo1112	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VARongo1112	VARongo1112
	VARXPNOGO4	BVARPNOGO	BVARPNOGO4	VARongo1112	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	VARongo1112	BVARPNOGO	VARongo12
	BVARPNOGO4	BVARPNOGO4	VARongo1112	BVARPNOGO4	VARPNOGO	VARongo1112	VARPNOGO	VARPNOGO	VARongo1112	BVARPNOGO4	BVARPNOGO4	BVARPNOGO
FIAT	BVARP	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	BVARP	VEPongo1112	PECM
	BVARPDNOGO4	VEPongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo12	VECPongo12	VADongo1112	VEPongo1112	VEPongo1112	PECM	VEPongo1112
	BVARPNOGO	VARPDNOGO	BVARPDNOGO4	BVARPNOGO	BVARPNOGO	BVARPNOGO	VARongo12	BVARP	VADongo1112	PECM	VADongo1112	BVARP
TOYOTA	BVARPNOGO4	BVARPDNOGO	BVARPDNOGO	VARongo1112	VEPongo1112	VECMXPNOGO4	VEPongo1112	BVARP	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPDNOGO	BVARPD	BVARPD	VEPongo1112	BVARP	VEPongo1112	VECMXPNOGO4	VARongo1112	BVARP	BVARP	BVARP	BVARP
	BVARP	BVARPDNOGO4	BVARPDNOGO4	BVARPNOGO4	VARongo1112	BVARP	VARongo1112	VEPongo1112	VADongo12	VARongo1112	VARongo1112	VARongo1112
JAGUAR	BVARPNOGO4	BVARPNOGO4	VARPNOGO	VARPNOGO	VARongo12	BVARPD	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARPNOGO	BVARPDNOGO4
	BVARP	BVARP	VARongo1112	VARongo12	VARPNOGO	BVARPDNOGO4	VARongo12	VARPNOGO	VARongo12	VARPNOGO	BVARPDNOGO4	BVARPD
	VARongo1112	VEPongo1112	VARPNOGO4	VARPNOGO4	BVARPNOGO4	VADongo1112	VARPNOGO	VARongo12	VARongo1112	VARongo12	BVARPD	BVARPDNOGO
KIA	AR12	BVARPDNOGO	BVARPDNOGO	AR12	AR12	BVARPD	AR12	AR12	AR12	AR12	AR12	VADongo1112
	VARPNOGO	BVARPNOGO	VARongo1112	VARongo1112	BVARPDNOGO	VADongo1112	BVARP	AR124	BVARPDNOGO	VARongo1112	BVARPDNOGO	BVARPDNOGO
	VARXPNOGO	AR12	BVARPNOGO	AR124	BVARP	BVARPDNOGO	VECMXPNOGO	VARongo1112	BVARPDNOGO4	AR124	AR124	BVARPDNOGO4
MITSUBISHI	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPDNOGO	BVARPD	BVARPD	BVARPD	BVARPDNOGO	VADongo1112	PECM	PECM	PECM	PECM	VECPongo12	PECM
	BVARPD	VADongo1112	BVARPDNOGO	BVARPDNOGO	VADongo1112	RW4	BVARPDNOGO	BVARPDNOGO	VECPongo12	VECPongo12	PECM	VECPongo12
SUZUKI	BVARPDNOGO	BVARPNOGO	VARongo1112	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4
	BVARPDNOGO4	BVARPDNOGO	BVARPNOGO4	BVARPNOGO	VARongo1112	BVARP	BVARP	BVARP	BVARP	BVARP	BVARP	BVARP
	BVARPD	BVARPNOGO4	BVARPNOGO	VARongo1112	BVARP	VARongo1112	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO

	Step 13	Step 14	Step 15	Step 16	Step 17	Step 18	Step 19	Step 20	Step 21	Step 22	Step 23	Step 24
BMW	VADongo1112	VADongo1112	AR12	AR12	AR124	AR124	VARPD	AR124	AR124	AR124	VECMPNOGO	AR12
	AR12	AR12	VADongo1112	AR124	AR12	AR12	AR124	AR12	AR12	AR12	VECMPNOGO4	AR124
	AR124	BVARPD	AR124	VADongo1112	VADongo1112	VARPD	AR12	VECMP	VECMPNOGO	VECMP	AR12	VARPNOGO
OPEL	VEPongo1112	VEPongo1112	VEPongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	VADongo1112	VADongo1112	VADongo12	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPD	BVARPDNOGO	VARongo1112	VADongo12	BVARPDNOGO	BVARPDNOGO	BVARPDNOGO	BVARPDNOGO	BVARPDNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO4
VOLKSWAGEN	BVARPNOGO	BVARPNOGO	BVARPNOGO	VECPongo12	VECPongo12	BVARPNOGO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VECPongo12
	BVARP	AR124	AR124	BVARPNOGO	BVARPNOGO	VECPongo12	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VADongo1112	VEPongo1112
	VEPongo1112	BVARP	AR12	VARongo1112	VARongo1112	VEPongo1112	BVARPDNOGO	BVARP	VADongo1112	VADongo1112	VECPongo12	VADongo1112
CITROEN	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARPNOGO	BVARPNOGO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARongo12	VEPongo1112	VEPongo1112
	BVARPNOGO4	VARongo12	BVARPNOGO	BVARPNOGO	VADongo1112	VADongo1112	VADongo1112	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO
FIAT	PECM	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	AR12	AR12	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	VEPongo1112	PECM	VADongo1112	VADongo1112	AR12	VEPongo1112	VARongo1112	VARPDNOGO	VARPDNOGO	VADongo12	VEPongo1112	VEPongo1112
	BVARP	VADongo1112	PECM	BVARPDNOGO4	BVARPDNOGO4	VARPDNOGO	BVARPDNOGO	VARXPDNOGO	VARXPDNOGO	VEPongo1112	VADongo12	RW
TOYOTA	BVARP	BVARPD	VARongo1112	BVARPD	VARongo1112	VEPongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARPD	BVARPDNOGO4	BVARPD	BVARPDNOGO	BVARPD	VARongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	VEPongo1112	BVARPDNOGO	BVARPDNOGO	VARongo1112	BVARPDNOGO	BVARPD	VARPNOGO4	VARPNOGO4	VARPNOGO4	VECMXPNOGO	BVARPD	VADongo12
JAGUAR	BVARPDNOGO4	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO
	VEPongo1112	VARPNOGO4	VARPNOGO4	VARPNOGO4	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARPNOGO4	VEPongo1112	VARPNOGO4	VARPNOGO4
	BVARPD	VEPongo1112	VEPongo1112	VEPongo1112	VARPNOGO4	VARPNOGO4	VARPNOGO4	VARPNOGO4	VEPongo1112	VARPNOGO4	VADongo12	VEPongo1112
KIA	VARongo1112	AR12	AR124	AR124	AR124	VEPongo1112	AR124	AR124	AR124	AR124	AR124	AR124
	AR12	AR124	AR12	AR12	AR12	AR124	AR12	AR12	AR12	AR12	VARongo12	VARongo12
	VADongo1112	VADongo1112	VARongo1112	VARongo1112	VARongo1112	AR12	VARongo1112	VARongo1112	VARongo12	VARongo12	AR12	AR12
MITSUBISHI	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARongo1112	VARongo1112	VEPongo1112	VARongo1112
	PECM	PECM	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VEPongo1112	VEPongo1112	VARongo1112	VEPongo1112
	VECPongo12	VARongo1112	VECPongo12	VECPongo12	VECPongo12	VECPongo12	VECPongo12	VECPongo12	VECPongo12	VECPongo12	VARongo12	VECPongo12
SUZUKI	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPD	BVARPD	BVARPD	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4
	VADongo1112	BVARPDNOGO	BVARPDNOGO	BVARPD	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	BVARPD	VADongo1112	VARPNOGO	BVARPNOGO	BVARPD

Table 17: Top three models in terms of MSPE for each forecasting horizon and each car brand.

BMW	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24
Total n. of models selected	22	14	16	18	20	12	18	16	17	10	5	4	6	7	5	5	8	18	17	22	21	21	19	16
Google models	10	5	8	8	10	5	9	8	8	8	2	1	2	2	1	2	4	7	9	11	10	10	9	6
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
OPEL	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	8	8	10	1	1	3	5	8	8	8	6	5	10	12	13	10	10	10	4	5	4	4	4	5
Google models	4	5	6	1	1	3	3	4	4	4	4	3	5	6	8	6	6	6	3	3	3	2	2	2
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	no	no	no	yes
VOLKSWAGEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	2	4	2	2	3	4	3	4	4	5	4	5	6	7	14	15	15	13	11	11	10	5	2	2
Google models	0	1	0	0	1	2	1	2	2	3	2	3	4	3	7	8	8	8	7	6	6	4	2	2
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
CITROEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	1	4	4	5	3	4	4	4	3	4	5	7	8	8	11	9	8	8	9	11	13	10	11	10
Google models	0	1	1	1	0	1	1	1	1	1	2	3	4	4	6	4	4	4	5	6	7	6	6	6
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
FIAT	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	5	4	8	6	9	10	12	10	13	12	12	8	9	9	8	9	8	3	4	2	3	1	2	4
Google models	2	2	4	3	6	8	8	7	7	6	6	4	5	5	5	5	3	1	1	1	1	1	2	2
Random Walk included?	no	no	no	no	no	no	no	no	no	no	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	yes
TOYOTA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	12	11	7	5	7	8	13	16	15	9	11	9	10	9	13	11	11	9	3	5	3	8	8	11
Google models	5	5	2	3	4	5	5	7	6	4	4	4	5	5	4	5	4	3	2	3	2	3	3	5
Random Walk included?	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	yes
JAGUAR	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	15	25	25	25	29	12	19	20	20	26	22	11	18	15	9	12	14	11	13	17	9	3	6	14
Google models	8	14	12	12	15	6	10	10	10	13	10	5	10	9	4	6	7	4	6	9	4	1	2	5
Random Walk included?	no	no	yes	no	no	no	no	no	no	no	no	yes	no	no	yes	no	no	no	no	no	no	no	no	yes
KIA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	2	11	10	8	13	12	3	6	11	11	12	12	15	13	14	15	13	14	14	11	11	9	7	7
Google models	0	2	4	2	5	5	1	2	4	3	4	5	7	5	6	6	5	6	7	4	4	4	2	2
Random Walk included?	no	yes	no	no	yes	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
MITSUBISHI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	12	7	12	12	12	15	14	1	4	4	6	6	5	5	4	4	4	4	2	2	2	3	3	4
Google models	6	4	5	6	7	8	8	1	4	4	5	5	5	5	4	4	4	4	2	2	2	3	1	3
Random Walk included?	no	no	yes	no	no	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
SUZUKI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	2	10	8	5	9	6	11	10	9	11	12	13	12	13	11	12	10	9	9	9	5	4	4	5
Google models	0	4	3	2	3	3	4	6	3	5	5	6	5	5	6	6	6	4	4	5	3	2	2	2
Random Walk included?	no	no	no	no	no	yes	yes	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	no	no

Table 18: Models included in the Model Confidence Set for each forecast horizon (from 1 step up to 24 steps ahead) and for each car brand.

C Robustness Checks: Nonlinear models

C.1 Seasonally Adjusted data

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12
BMW	AAR(6)log	SETAR(1)log	AR12	AR12	SETAR(8)log	AAR(7)dlog	AAR(6)dlog	AR12	AAR(5)dlog	SETAR(1)log	SETAR(1)log	SETAR(1)log
	AAR(8)log	SETAR(2)log	SETAR(6)dlog	SETAR(1)log	AR12	AAR(6)dlog	SETAR(3)dlog	SETAR(1)log	AAR(6)dlog	AAR(5)dlog	AR12	AAR(1)log
	AAR(7)log	AR12	SETAR(5)dlog	SETAR(6)log	SETAR(1)log	AAR(9)dlog	AR12	SETAR(2)log	AR12	SETAR(2)log	AAR(1)log	LSTAR(1)log
OPEL	LSTAR(3)dlog	VADongo12	VADongo12	VECongo1112	VECongo1112	VECongo1112	VECM	VECongo1112	VECongo1112	VECMXNOGO	VECongo1112	SETAR(7)dlog
	LSTAR(2)dlog	VECongo1112	VECongo1112	VADongo12	VECM	VECM	VECMX	VARongo1112	VARongo1112	VECMNOGO	VECMXNOGO	AAR(8)dlog
	LSTAR(4)dlog	VARongo12	VARongo12	VECongo12	VECMX	VECMX	VECongo1112	VECMXNOGO	VECMXNOGO	VECongo1112	VECMNOGO	VECongo1112
VOLKSWAGEN	BVARDNOGO	BVARDNOGO	VARNOGO	VARDNOGO	AR124	AR124	VARongo1112	AR124	AR124	AR124	AR124	AR124
	BVARDNOGO4	BVARD	VARXNOGO	VARXDNOGO	VARongo1112	VARongo1112	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	AAR(1)dlog	BVARDNOGO4	VARDNOGO	VARNOGO	VECongo1112	AR12	AR12	AR12	AR12	AR12	AR12	AR12
CITROEN	VARDNOGO4	VECMXNOGO4	VECMXNOGO4	VARD	VARD	LSTAR(3)log	LSTAR(3)log	LSTAR(3)log	SETAR(4)log	SETAR(4)log	SETAR(4)log	SETAR(4)log
	VARXDNOGO4	VECMNOGO4	VECMNOGO4	VARXD	VARXD	SETAR(4)log	SETAR(4)log	SETAR(4)log	LSTAR(3)log	SETAR(5)log	LSTAR(3)log	LSTAR(3)log
	VECMXNOGO4	BVARDNOGO	VARD	VARDNOGO	VARDNOGO	SETAR(5)log	SETAR(5)log	SETAR(5)log	SETAR(5)log	LSTAR(3)log	SETAR(3)log	SETAR(3)log
FIAT	VECongo1112	AAR(1)log	AAR(1)log	SETAR(4)log	SETAR(2)log	SETAR(2)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	VARongo1112	VADongo1112	SETAR(4)log	AAR(1)log	AAR(1)log	AAR(1)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log
	VADongo1112	VECongo1112	SETAR(3)log	SETAR(5)log	SETAR(4)log	SETAR(4)log	LSTAR(2)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log
TOYOTA	AAR(1)log	AAR(1)log	VARNOGO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	BVAR	SETAR(6)dlog	SETAR(6)dlog	BVARNOGO4
	SETAR(2)log	VECMXNOGO	VARXNOGO	VARNOGO	VARNOGO	BVAR	BVAR	BVAR	BVARNOGO4	BVARNOGO4	BVARNOGO4	SETAR(7)dlog
	SETAR(1)log	VECMNOGO	VECongo1112	VARXNOGO	VARXNOGO	VECMXNOGO	BVARNOGO4	BVARNOGO4	SETAR(6)dlog	BVAR	BVAR	SETAR(5)dlog
JAGUAR	AAR(4)log	SETAR(4)dlog	AAR(3)log	AAR(3)log	AAR(4)log	SETAR(5)log	VECongo1112	SETAR(3)log	LSTAR(4)log	LSTAR(4)log	LSTAR(4)log	SETAR(3)log
	AAR(3)log	AAR(3)log	AAR(4)log	LSTAR(3)log	SETAR(5)log	LSTAR(4)log	SETAR(5)log	LSTAR(4)log	LSTAR(5)log	SETAR(3)log	SETAR(3)log	LSTAR(5)log
	AAR(5)log	SETAR(3)dlog	LSTAR(1)dlog	AAR(4)log	SETAR(3)log	SETAR(4)log	AAR(4)log	SETAR(4)log	SETAR(3)log	AAR(4)log	AAR(4)log	AR12
KIA	BVARDNOGO	AAR(7)log	AAR(7)log	AAR(7)log	AAR(7)log	AAR(4)log	SETAR(3)log	SETAR(3)log	AAR(7)log	SETAR(3)log	SETAR(3)log	SETAR(3)log
	BVARDNOGO4	SETAR(9)log	AAR(4)log	AAR(4)log	AAR(4)log	AAR(2)log	AAR(4)log	AAR(4)log	SETAR(3)log	LSTAR(1)log	SETAR(1)log	LSTAR(1)log
	BVARD	SETAR(1)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(7)log	AAR(2)log	LSTAR(1)log	LSTAR(1)log	AAR(7)log	LSTAR(1)log	AAR(4)log
MITSUBISHI	LSTAR(2)log	LSTAR(2)log	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo12	VECongo12	VECongo12
	BVARDNOGO	VECongo1112	AAR(4)dlog	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VECongo1112	VARongo1112	VECongo1112
	LSTAR(5)log	AAR(2)dlog	AAR(5)dlog	AAR(4)log	SETAR(1)log	SETAR(1)log	VARongo12	VARongo12	VARongo12	VARongo1112	VARongo12	VARongo1112
SUZUKI	RW	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log
	RW4	SETAR(2)log	SETAR(5)log	SETAR(5)log	VECMXNOGO	SETAR(2)log	SETAR(2)log	AAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log
	VARongo1112	VARongo1112	AAR(4)log	VECMXNOGO	VECMNOGO	LSTAR(1)log	AAR(3)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log

	Step 13	Step 14	Step 15	Step 16	Step 17	Step 18	Step 19	Step 20	Step 21	Step 22	Step 23	Step 24
BMW	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	AR12	AR12	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	SETAR(1)log	VARongo1112	AR12	AR12	AR12	AR12	AR124	AR124	AR124	AR124	AR124	AR124
OPEL	SETAR(7)dlog	AAR(8)dlog	AAR(8)dlog	AAR(8)dlog	SETAR(7)dlog	SETAR(7)dlog	AAR(8)dlog	AAR(8)dlog	SETAR(7)dlog	SETAR(7)dlog	SETAR(8)dlog	SETAR(8)dlog
	AAR(8)dlog	SETAR(7)dlog	SETAR(7)dlog	SETAR(7)dlog	AAR(8)dlog	AAR(8)dlog	SETAR(7)dlog	SETAR(7)dlog	SETAR(8)dlog	SETAR(8)dlog	SETAR(7)dlog	VECongo1112
	SETAR(8)dlog	SETAR(8)dlog	SETAR(8)dlog	SETAR(8)dlog	AAR(7)dlog	AAR(7)dlog	AAR(3)dlog	SETAR(5)dlog	AAR(8)dlog	AAR(8)dlog	VECongo1112	SETAR(7)dlog
VOLKSWAGEN	AR124	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AR124	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	AR12	VARongo1112	AR124	AR124	AR124	AR124	AR124	AR12	AR124	AR124	AR124	AR124
	VARongo1112	AR12	AR12	AR12	AR12	AR12	AR12	VARongo1112	AR12	AR12	AR12	AR12
CITROEN	SETAR(4)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	LSTAR(3)log	BVARNOGO	LSTAR(3)log	SETAR(4)log	LSTAR(3)log	SETAR(4)log	LSTAR(3)log	SETAR(3)log	SETAR(4)log	SETAR(3)log	SETAR(4)log	SETAR(4)log
	AAR(1)log	SETAR(4)log	SETAR(4)log	LSTAR(3)log	SETAR(4)log	SETAR(5)log	SETAR(4)log	LSTAR(3)log	LSTAR(3)log	LSTAR(3)log	LSTAR(3)log	SETAR(3)log
FIAT	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	SETAR(2)log	SETAR(2)log	LSTAR(2)log	LSTAR(2)log	LSTAR(2)log	LSTAR(2)log	SETAR(2)log	AR124	AR124	AR124	AR124	AR124
	SETAR(1)log	LSTAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	AR124	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log
TOYOTA	SETAR(5)dlog	SETAR(5)dlog	SETAR(6)dlog	SETAR(6)dlog	SETAR(3)dlog	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	SETAR(4)dlog	SETAR(4)dlog	SETAR(5)dlog	SETAR(3)dlog	SETAR(5)dlog	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog	SETAR(6)dlog	SETAR(6)dlog	SETAR(7)dlog	SETAR(6)dlog
	SETAR(3)dlog	SETAR(3)dlog	SETAR(4)dlog	SETAR(5)dlog	SETAR(4)dlog	SETAR(6)dlog	SETAR(7)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog
JAGUAR	AR12	AR12	AR12	VECongo1112	VECongo1112	AR12	SETAR(1)log	VECongo1112	VARongo1112	VARongo12	VECongo1112	AR12
	SETAR(3)log	VECongo1112	SETAR(3)log	AR12	AR12	VECongo1112	VECongo1112	SETAR(1)log	VARongo12	AR12	AR12	VECongo1112
	AR124	SETAR(1)log	VECongo1112	SETAR(1)log	VARongo1112	VARongo1112	AR12	AR12	AR12	VARongo1112	VARongo1112	AR124
KIA	LSTAR(1)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(6)log	SETAR(6)log	SETAR(6)log
	SETAR(3)log	SETAR(3)log	AAR(8)log	SETAR(6)log	SETAR(5)log	SETAR(7)log	SETAR(7)log	SETAR(6)log	SETAR(6)log	SETAR(8)log	LSTAR(1)log	LSTAR(1)log
	AAR(4)log	LSTAR(1)log	SETAR(1)log	SETAR(5)log	SETAR(8)log	LSTAR(1)log	SETAR(6)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	SETAR(8)log	SETAR(8)log
MITSUBISHI	VECongo1112	VECongo1112	VARongo1112	VARongo1112	VARongo1112	VECongo1112	VECongo1112	VECMNOGO4	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	VARongo1112	VARongo1112	VECongo1112	VECongo1112	VECongo1112	VARongo1112	VECMNOGO4	VARongo12	VECMNOGO4	VECMNOGO4	VECMNOGO4	SETAR(1)log
	VECongo12	VARongo12	SETAR(1)log	VARongo12	VARongo12	VARongo12	VARongo1112	VECongo1112	VARongo1112	VARongo1112	VARongo12	VECMNOGO4
SUZUKI	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	SETAR(2)log
	SETAR(3)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	AAR(1)log
	AAR(3)log	SETAR(3)log	AAR(2)log	SETAR(3)log	AAR(2)log	SETAR(3)log	AAR(2)log	AAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log

Table 19: Top three models in terms of MSPE for each forecasting horizon and each car brand.

BMW	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	49	72	65	75	71	71	73	72	73	73	63	57	37	37	37	36	8	3	7	3	6	6	8	6
Google models	6	13	12	13	13	13	13	12	12	12	9	6	1	4	4	4	2	1	2	1	2	1	1	1
Nonlinear models	36	42	42	43	43	42	42	43	43	43	41	38	31	29	29	27	4	1	3	1	2	2	4	2
Random Walk included?	no	yes	no	yes	no	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
OPEL	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	16	4	8	8	2	2	1	4	5	1	2	10	31	40	31	31	40	34	39	31	18	13	7	6
Google models	1	3	5	6	2	2	1	3	3	0	1	2	4	4	3	3	4	4	5	3	2	2	1	1
Nonlinear models	12	0	2	0	0	0	0	0	0	0	0	5	22	31	24	24	30	25	27	22	15	10	5	4
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	yes	yes	yes	yes	no	no	no	no
VOLKSWAGEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	7	4	1	1	3	4	6	1	3	3	4	6	4	5	4	4	4	4	4	5	4	4	4	4
Google models	2	1	0	0	1	1	2	0	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1
Nonlinear models	1	0	0	0	0	0	1	0	1	1	2	3	1	2	1	1	1	1	1	1	1	1	1	1
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
CITROEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	1	1	1	80	81	4	10	12	13	14	12	8	13	12	13	21	15	34	27	8	10	8	7	6
Google models	0	0	0	14	14	0	0	1	1	2	1	1	2	2	3	3	2	6	3	1	1	1	1	0
Nonlinear models	0	0	0	46	47	3	7	9	10	10	9	5	9	8	8	15	12	23	20	7	9	7	6	6
Random Walk included?	no	no	no	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
FIAT	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	14	16	32	34	40	42	45	35	36	25	16	1	1	1	1	1	2	2	2	2	2	2	2	2
Google models	6	2	2	3	4	4	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Nonlinear models	4	12	25	25	29	30	31	24	26	18	10	1	1	1	1	1	1	1	1	1	1	1	1	1
Random Walk included?	yes	no	no	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no
TOYOTA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	3	2	1	2	2	5	5	8	9	13	23	22	25	34	30	36	21	30	22	31	23	22	22	7
Google models	0	0	0	1	1	2	2	2	2	2	2	2	3	3	2	5	2	3	2	2	2	2	2	2
Nonlinear models	2	1	0	0	0	1	1	3	5	9	17	16	19	24	21	25	16	22	16	23	18	17	18	5
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no	no	no
JAGUAR	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	65	69	64	74	70	76	71	72	76	74	75	59	44	46	35	33	25	32	33	27	29	29	32	23
Google models	6	7	8	9	8	10	8	8	9	10	10	7	9	7	7	6	4	5	6	5	5	4	5	4
Nonlinear models	49	50	46	53	53	54	53	53	53	51	53	43	27	34	24	25	18	21	24	20	22	23	22	15
Random Walk included?	no	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no
KIA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	86	86	85	85	84	85	85	85	85	85	64	52	28	24	79	69	36	24	23	17	13	10	12	10
Google models	14	14	14	14	14	14	14	14	14	14	10	9	2	3	13	9	6	2	2	2	2	1	1	0
Nonlinear models	52	52	51	51	51	51	51	51	51	51	46	37	21	19	50	47	24	20	20	14	10	9	11	10
Random Walk included?	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	no	no	no	no	no	no	no	no
MITSUBISHI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	98	98	98	98	98	98	98	98	98	97	82	65	39	56	40	34	24	6	6	13	12	6	14	17
Google models	14	14	14	14	14	14	14	14	14	14	11	10	6	8	5	6	5	3	3	5	4	3	5	5
Nonlinear models	64	64	64	64	64	64	64	64	64	63	55	44	27	39	28	22	15	1	1	5	6	1	6	8
Random Walk included?	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
SUZUKI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	11	18	16	4	2	13	27	23	20	16	12	8	10	8	8	7	7	8	7	7	7	7	6	5
Google models	4	4	3	0	0	2	4	4	3	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Nonlinear models	2	11	11	3	1	10	20	15	13	13	10	8	10	8	8	7	7	8	7	7	7	7	6	5
Random Walk included?	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no

Table 20: Models included in the Model Confidence Set for each forecast horizon (from 1 step up to 24 steps ahead) and for each car brand.

C.2 Raw data

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12
BMW	BVARPNOGO4	BVARPD	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	AR12	AR124
	BVARP	BVARP	BVARPDNOGO	BVARPDNOGO	AR12	BVARPDNOGO	BVARPD	BVARPD	BVARPDNOGO	AR12	AR124	AR12
	BVARPNOGO	BVARPDNOGO	BVARPDNOGO4	BVARPD	BVARPDNOGO	AR12	BVARPDNOGO4	BVARPDNOGO4	AR12	AR124	VADongo1112	VADongo1112
OPEL	BVARPDNOGO	BVARPNOGO	VARongo1112	VARP	VARP	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPNOGO	VEPongo1112	VEPongo1112	VARXP	VARXP	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VADongo1112	VADongo1112
	BVARPDNOGO4	BVARPDNOGO	BVARPD	VEPongo1112	VEPongo1112	VARP	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VARongo1112	RW
VOLKSWAGEN	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO
	VECMXPNOGO	BVARP	VARPNOGO	VARPNOGO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	BVARP	SETAR(7)log	SETAR(4)log	SETAR(6)log
	VECMPNOGO	BVARPDNOGO	VARXPNOGO	VARXPNOGO	VARPNOGO	VADongo1112	VARPNOGO	BVARP	VEPongo1112	SETAR(3)log	SETAR(3)log	SETAR(2)log
CITROEN	VARPNOGO4	VARongo1112	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VARongo1112	VARongo1112
	VARXPNOGO4	BVARPNOGO	BVARPNOGO4	VARongo1112	BVARPNOGO4	BVARPNOGO4	BVARPNOGO4	AAR(3)log	AAR(3)log	VARongo1112	AAR(3)log	VARongo12
	BVARPNOGO4	BVARPNOGO4	VARongo1112	BVARPNOGO4	VARPNOGO	VARongo1112	VARPNOGO	BVARPNOGO4	BVARPNOGO4	AAR(3)log	BVARPNOGO	AAR(3)log
FIAT	BVARP	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AAR(6)log	AAR(6)log	AAR(6)log	SETAR(1)log
	BVARPDNOGO4	VEPongo1112	VADongo1112	VADongo1112	LSTAR(2)log	LSTAR(2)log	SETAR(1)log	AAR(6)log	AAR(5)log	AAR(5)log	SETAR(1)log	SETAR(4)log
	BVARPNOGO	VARPDNOGO	BVARPDNOGO4	LSTAR(2)log	LSTAR(3)log	AAR(1)log	AAR(1)log	SETAR(1)log	AAR(3)log	SETAR(4)log	SETAR(4)log	AAR(1)log
TOYOTA	BVARPNOGO4	BVARPDNOGO	BVARPDNOGO	VARongo1112	VEPongo1112	VECMXPNOGO4	VEPongo1112	BVARP	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPDNOGO	BVARPD	BVARPD	VEPongo1112	BVARP	VEPongo1112	VECMXPNOGO4	VARongo1112	BVARP	BVARP	BVARP	BVARP
	BVARP	BVARPDNOGO4	BVARPDNOGO4	BVARPNOGO4	VARongo1112	BVARP	VARongo1112	VEPongo1112	VADongo12	VARongo1112	VARongo1112	VARongo1112
JAGUAR	BVARPNOGO4	BVARPNOGO4	VARPNOGO	VARPNOGO	VARongo12	BVARPD	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARPNOGO	BVARPDNOGO4
	BVARP	BVARP	VARongo1112	VARongo12	VARPNOGO	BVARPDNOGO4	VARongo12	VARPNOGO	VARongo12	VARPNOGO	BVARPDNOGO4	BVARPD
	VARongo1112	VEPongo1112	VARPNOGO4	VARPNOGO4	BVARPNOGO4	VADongo1112	VARPNOGO	VARongo12	VARongo1112	VARongo12	BVARPD	BVARPDNOGO
KIA	AAR(4)log	AAR(4)log	AAR(7)log	AR12	AAR(9)log	BVARPD	AAR(7)log	AAR(7)log	AAR(7)log	AAR(9)log	AAR(9)log	AAR(9)log
	AAR(6)log	BVARPDNOGO	BVARPDNOGO	AAR(9)log	AR12	VADongo1112	AAR(9)log	AAR(9)log	AAR(9)log	AAR(7)log	AAR(8)log	AAR(7)log
	AAR(2)log	AAR(6)log	VARongo1112	AAR(7)log	AAR(7)log	BVARPDNOGO	AR12	SETAR(5)log	AR12	AR12	AAR(7)log	VADongo1112
MITSUBISHI	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPDNOGO	AAR(1)log	AAR(1)log	LSTAR(3)log	LSTAR(2)log	LSTAR(3)log	LSTAR(2)log	LSTAR(2)log	PECM	LSTAR(2)log	LSTAR(2)log	AAR(1)log
	BVARPD	AAR(2)log	BVARPD	LSTAR(4)log	LSTAR(3)log	AAR(2)log	LSTAR(4)log	AAR(1)log	LSTAR(2)log	AAR(1)log	LSTAR(3)log	AAR(2)log
SUZUKI	BVARPDNOGO	BVARPNOGO	SETAR(2)log	AAR(1)log	SETAR(2)log	SETAR(2)log	SETAR(3)log	AAR(2)log	SETAR(2)log	SETAR(3)log	AAR(4)log	AAR(4)log
	BVARPDNOGO4	BVARPDNOGO	AAR(1)log	SETAR(2)log	AAR(1)log	SETAR(3)log	SETAR(2)log	SETAR(3)log	AAR(2)log	AAR(2)log	SETAR(6)log	SETAR(3)log
	BVARPD	BVARPNOGO4	VARongo1112	AAR(3)log	AAR(3)log	AAR(3)log	AAR(2)log	SETAR(2)log	AAR(3)log	SETAR(2)log	AAR(3)log	AAR(3)log

	Step 13	Step 14	Step 15	Step 16	Step 17	Step 18	Step 19	Step 20	Step 21	Step 22	Step 23	Step 24
BMW	VADongo1112	LSTAR(1)log0	LSTAR(1)log0	LSTAR(1)log0	LSTAR(8)log	LSTAR(9)log	LSTAR(1)log1	LSTAR(9)log	SETAR(7)log	LSTAR(1)log1	LSTAR(1)log1	LSTAR(8)log
	LSTAR(9)log	VADongo1112	LSTAR(9)log	AR12	SETAR(1)log	SETAR(9)log	LSTAR(9)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	LSTAR(7)log	LSTAR(1)log1
	LSTAR(1)log0	LSTAR(9)log	AR12	LSTAR(8)log	LSTAR(1)log1	SETAR(1)log	LSTAR(1)log0	LSTAR(8)log	LSTAR(1)log0	SETAR(6)log	LSTAR(8)log	SETAR(1)log
OPEL	VEPongo1112	VEPongo1112	VEPongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	VADongo1112	VADongo1112	VADongo12	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	BVARPD	BVARPDNOGO	VARongo1112	VADongo12	BVARPDNOGO	BVARPDNOGO	BVARPDNOGO	BVARPDNOGO	BVARPDNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO4
VOLKSWAGEN	BVARPNOGO	BVARPNOGO	BVARPNOGO	VECPongo12	VECPongo12	BVARPNOGO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VECPongo12
	SETAR(2)log	AR124	AR124	BVARPNOGO	BVARPNOGO	VECPongo12	BVARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VADongo1112	VEPongo1112
	SETAR(3)log	BVARP	SETAR(3)log	VARongo1112	VARongo1112	VEPongo1112	BVARPDNOGO	BVARP	VADongo1112	VADongo1112	VECPongo12	VADongo1112
CITROEN	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log
	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	VARongo1112	VARongo1112	VARongo1112	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(2)dlog	SETAR(2)dlog
	BVARPNOGO	BVARPNOGO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	SETAR(1)log	VARongo1112	SETAR(2)dlog	SETAR(2)dlog	LSTAR(2)dlog	LSTAR(2)dlog
FIAT	SETAR(1)log	LSTAR(4)log	LSTAR(2)log	LSTAR(2)log	LSTAR(2)log	LSTAR(2)log	AAR(3)log	AAR(6)log	AAR(5)log	LSTAR(1)log1	VARongo1112	LSTAR(5)log
	LSTAR(4)log	SETAR(1)log	LSTAR(4)log	SETAR(3)log	LSTAR(3)log	LSTAR(3)log	LSTAR(2)log	SETAR(2)log	AAR(6)log	SETAR(1)log	LSTAR(5)log	AAR(3)log
	SETAR(4)log	LSTAR(2)log	SETAR(3)log	LSTAR(3)log	SETAR(4)log	AAR(3)log	LSTAR(3)log	AAR(5)log	SETAR(2)log	LSTAR(5)log	AAR(4)log	AAR(4)log
TOYOTA	BVARP	BVARPD	VARongo1112	BVARPD	VARongo1112	VEPongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARPD	BVARPDNOGO4	BVARPD	BVARPDNOGO	BVARPD	VARongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112
	VEPongo1112	BVARPDNOGO	BVARPDNOGO	VARongo1112	BVARPDNOGO	SETAR(8)log	VARPNOGO4	SETAR(3)log	SETAR(6)log	SETAR(3)log	SETAR(3)log	VADongo12
JAGUAR	BVARPDNOGO4	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO
	VEPongo1112	VARPNOGO4	VARPNOGO4	VARPNOGO4	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	AAR(4)log	LSTAR(1)log0	VARPNOGO4	VARPNOGO4
	BVARPD	VEPongo1112	VEPongo1112	VEPongo1112	VARPNOGO4	VARPNOGO4	VARPNOGO4	LSTAR(1)log0	AAR(3)log	AAR(3)log	LSTAR(1)log0	VEPongo1112
KIA	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log	AAR(9)log
	SETAR(5)log	AAR(8)log	AAR(7)log	AAR(7)log	AAR(7)log	AAR(7)log	AAR(7)log	AAR(7)log	AAR(7)log	AAR(7)log	AR124	AAR(7)log
	AAR(6)log	SETAR(5)log	AAR(6)log	SETAR(5)log	SETAR(5)log	AAR(8)log	SETAR(7)log	SETAR(5)log	AR124	AR124	AAR(7)log	AAR(8)log
MITSUBISHI	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARongo1112	VARongo1112	VEPongo1112	VARongo1112
	AAR(1)log	AAR(1)log	AAR(1)log	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VEPongo1112	VEPongo1112	VARongo1112	VEPongo1112
	AAR(2)log	PECM	LSTAR(2)log	AAR(1)log	AAR(1)log	VECPongo12	VECPongo12	VECPongo12	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
SUZUKI	AAR(2)log	SETAR(7)log	SETAR(6)log	SETAR(3)log	AAR(4)log	AAR(2)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	AAR(2)log	AAR(2)log	AAR(2)log
	SETAR(4)log	AAR(2)log	AAR(2)log	SETAR(4)log	AAR(2)log	SETAR(6)log	AAR(2)log	SETAR(5)log	AAR(2)log	AAR(1)log	AAR(1)log	AAR(1)log
	SETAR(3)log	SETAR(2)log	AAR(4)log	AAR(4)log	SETAR(6)log	AAR(4)log	AAR(4)log	AAR(2)log	AAR(4)log	SETAR(6)log	SETAR(2)log	SETAR(7)log

Table 21: Top three models in terms of MSPE for each forecasting horizon and each car brand.

BMW	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	19	12	14	18	23	12	18	20	30	14	9	4	11	21	22	14	38	48	57	60	55	57	65	59
Google models	7	4	6	8	10	5	9	9	10	3	3	1	2	4	4	2	7	9	13	13	12	13	13	10
Nonlinear models	2	0	0	0	3	0	0	3	10	2	3	0	5	12	14	9	22	30	33	34	32	32	39	36
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes
OPEL	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	8	10	10	1	1	3	5	8	12	6	6	5	13	18	23	15	16	11	4	5	6	4	4	4
Google models	4	6	6	1	1	3	3	4	5	4	4	3	6	7	9	7	6	6	3	3	3	2	2	2
Nonlinear models	0	0	0	0	0	0	0	0	3	0	0	0	3	6	9	5	6	2	0	0	1	0	0	0
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	no	no	no	yes
VOLKSWAGEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	2	4	2	2	3	4	3	4	5	12	13	16	16	16	21	24	28	20	11	17	15	5	2	2
Google models	0	1	0	0	1	2	1	2	2	3	4	4	4	5	7	8	9	7	6	6	6	4	2	2
Nonlinear models	0	0	0	0	0	0	0	0	1	5	6	7	7	5	7	9	11	7	1	7	4	0	0	0
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	no	no	no	no
CITROEN	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	1	4	4	5	3	4	5	5	5	7	8	10	19	20	32	40	42	47	48	40	48	49	50	41
Google models	0	1	1	1	0	1	1	1	1	2	2	3	5	5	6	5	5	5	5	6	7	6	7	3
Nonlinear models	0	0	0	0	0	0	1	1	1	2	3	4	11	12	21	30	32	36	38	29	38	38	38	35
Random Walk included?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no
FIAT	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	5	4	11	10	20	23	24	27	29	26	31	24	27	17	11	9	18	25	25	22	27	24	25	24
Google models	2	2	5	2	6	7	7	8	7	4	5	4	5	3	2	0	2	2	2	1	2	2	3	2
Nonlinear models	0	0	2	5	11	14	15	16	19	19	21	17	20	14	9	9	15	18	18	20	21	20	20	20
Random Walk included?	no	no	no	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
TOYOTA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	12	11	7	5	7	11	19	37	50	13	20	14	10	7	25	10	23	21	3	6	3	13	11	23
Google models	5	5	2	3	4	3	5	7	7	4	5	4	5	4	5	4	5	4	2	3	2	3	3	6
Nonlinear models	0	0	0	0	0	4	6	21	34	4	9	5	0	0	11	2	10	10	0	2	0	5	4	11
Random Walk included?	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	yes
JAGUAR	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	15	32	70	71	35	12	26	73	74	56	35	26	24	51	60	51	51	54	52	49	55	52	51	50
Google models	7	14	14	15	15	6	10	14	13	15	10	9	10	13	13	12	13	12	14	11	11	12	10	11
Nonlinear models	2	6	41	41	6	0	6	47	47	27	14	5	5	27	33	27	27	29	27	27	31	29	27	27
Random Walk included?	no	no	no	no	no	no	no	no	yes	no	no	yes	no	no	yes	no	no	no	no	no	yes	no	yes	yes
KIA	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	11	32	21	12	29	18	6	20	17	35	44	38	44	43	40	8	12	21	16	8	16	3	2	2
Google models	0	1	4	2	5	5	1	2	1	2	4	4	6	6	6	0	1	5	3	1	3	0	0	0
Nonlinear models	9	23	11	4	17	6	3	14	12	26	32	27	30	29	26	7	9	12	10	6	9	2	1	2
Random Walk included?	no	yes	no	no	no	no	no	no	no	yes	yes	yes	yes	yes	yes	no	no	no	yes	no	yes	no	no	no
MITSUBISHI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	22	21	50	74	80	86	25	66	69	72	82	83	56	63	61	26	21	21	28	43	35	26	40	44
Google models	6	6	7	10	11	12	5	8	9	10	12	12	8	10	9	4	5	6	7	7	6	7	6	6
Nonlinear models	10	10	36	55	57	57	17	53	54	55	56	55	43	46	44	21	16	14	20	34	28	19	33	37
Random Walk included?	no	no	yes	no	yes	yes	no	no	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
SUZUKI	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S19**	**S20**	**S21**	**S22**	**S23**	**S24**
Total n. of models selected	34	25	24	21	19	21	13	10	8	17	10	12	20	18	7	15	16	9	12	13	17	13	7	11
Google models	4	3	3	3	2	3	1	0	0	1	1	1	3	2	1	1	1	1	1	1	1	1	0	0
Nonlinear models	22	17	16	13	14	16	10	10	8	14	9	10	14	14	6	13	14	8	11	12	13	12	7	11
Random Walk included?	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no

Table 22: Models included in the Model Confidence Set for each forecast horizon (from 1 step up to 24 steps ahead) and for each car brand.

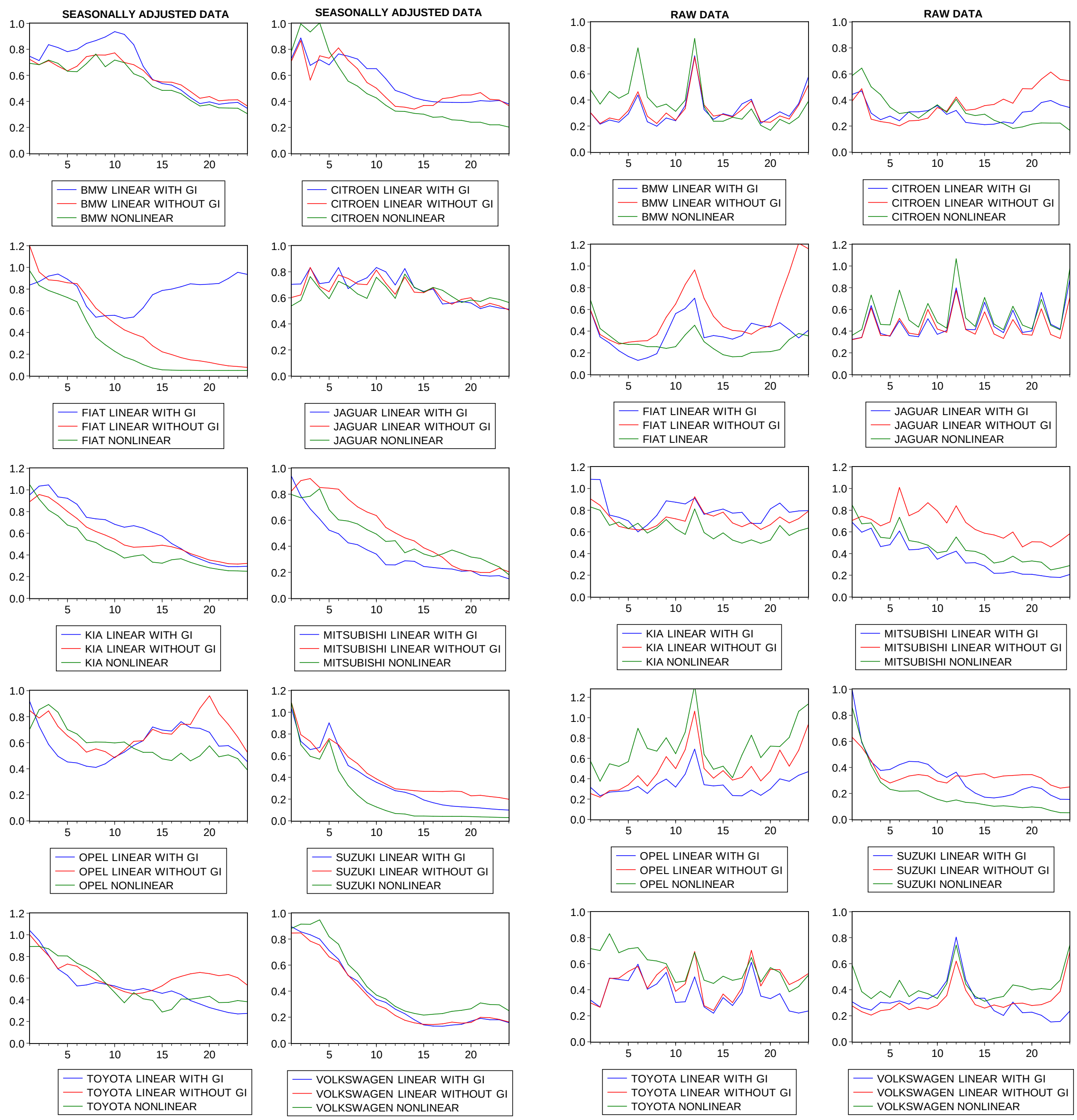

Figure 5: Ratios of the MSPEs of the best models with and without Google data with those of the Random Walk model, together with the ratios of the MSPEs of the best nonlinear models and the Random Walk model across all forecasting horizons. The first two columns show results for seasonally-adjusted data, and the last two for raw data.

C.3 List of Models included in the Model Confidence Set. Forecast horizon: 24 steps ahead. Seasonally Adjusted data

SEASONALLY ADJUSTED DATA									
BmwBMW	**Citroen**	**Fiat**	**Jaguar**	**Kia**	**Mitsubishi**	**Opel**	**Suzuki**	**Toyota**	**Volkswagen**
AAR(1)log	AAR(1)log	AAR(1)log	AR12	SETAR(6)log	VECongo1112	SETAR(8)dlog	SETAR(2)log	VECongo1112	VARongo1112
VARongo1112	SETAR(4)log	AR124	VECongo1112	LSTAR(1)log	SETAR(1)log	VECongo1112	AAR(1)log	SETAR(6)dlog	AR124
AR124	SETAR(3)log		AR124	SETAR(8)log	VECMNOGO4	SETAR(7)dlog	SETAR(3)log	SETAR(1)dlog	AR12
AR12	LSTAR(3)log		VARongo1112	SETAR(5)log	VARongo12	SETAR(5)dlog	AAR(3)log	SETAR(7)dlog	AAR(1)log
AAR(4)log	SETAR(5)log		AAR(2)log	SETAR(1)log	VARongo1112	AAR(8)dlog	SETAR(4)log	SETAR(3)dlog	
BVARNOGO	AAR(2)log		SETAR(1)log	SETAR(7)log	LSTAR(1)log0	VECMNOGO		SETAR(4)dlog	
			AAR(6)log	SETAR(1)log	SETAR(1)log				
			AAR(1)log	SETAR(9)log	AR12				
			LSTAR(1)log	SETAR(1)log	VECongo12				
			AAR(7)log	SETAR(3)log	VADongo12				
			AAR(3)log		LSTAR(1)log1				
			SETAR(3)log		LSTAR(1)log2				
			AAR(5)log		AR124				
			AAR(4)log		SETAR(4)log				
			VECongo12		AAR(5)log				
			LSTAR(4)log		AAR(3)dlog				
			LSTAR(5)log		VECMNOGO				
			LSTAR(3)log						
			SETAR(4)log						
			VARongo12						
			SETAR(1)log						
			VECMNOGO						
			VECMNOGO4						

Table 23: List of Models included in the Model Confidence Set for each car brand. Forecast horizon: 24 steps ahead.

C.4 List of Models included in the Model Confidence Set. Forecast horizon: 24 steps ahead. Raw data

RAW DATA									
BmwBMW	**Citroen**	**Fiat**	**Jaguar**	**Kia**	**Mitsubishi**	**Opel**	**Suzuki**	**Toyota**	**Volkswagen**
LSTAR(8)log	AAR(3)log	LSTAR(5)log	VARPNOGO	AAR(9)log	VARongo1112	VARongo1112	AAR(2)log	VARongo1112	VECPongo12
LSTAR(1)log1	SETAR(2)dlog	AAR(3)log	VARPNOGO4	AAR(8)log	VEPongo1112	VEPongo1112	AAR(1)log	VEPongo1112	VEPongo1112
SETAR(1)log	LSTAR(2)dlog	AAR(4)log	VEPongo1112		AAR(1)log	VECMPNOGO	SETAR(7)log	VADongo12	
SETAR(3)log	SETAR(1)log	SETAR(4)log	BVARPDNOGO4		AAR(2)log	RW	SETAR(3)log	SETAR(6)log	
LSTAR(1)log0	LSTAR(1)dlog	VARongo1112	BVARPD		LSTAR(5)log		SETAR(2)log	SETAR(8)log	
SETAR(1)log	SETAR(3)log	LSTAR(4)log	VADongo12		LSTAR(4)log		SETAR(4)log	SETAR(3)log	
LSTAR(7)log	LSTAR(4)dlog	SETAR(1)log	VARXPNOGO		VECPongo12		AAR(3)log	VECMPNOGO	
SETAR(7)log	SETAR(3)dlog	AAR(1)log	BVARPDNOGO		LSTAR(3)log		AAR(4)log	SETAR(5)log	
LSTAR(6)log	SETAR(4)dlog	LSTAR(3)log	VADongo1112		SETAR(5)log		SETAR(8)log	SETAR(4)log	
AR12	VARongo1112	SETAR(5)log	VECPongo12		VARongo12		SETAR(6)log	BVARPD	
SETAR(5)log	LSTAR(3)dlog	AAR(2)log	LSTAR(1)log0		AAR(3)log		AAR(7)dlog	SETAR(9)log	
AAR(1)log	SETAR(6)log	LSTAR(2)log	LSTAR(9)log		SETAR(4)log			BVARPDNOGO4	
SETAR(6)log	SETAR(4)log	AAR(6)log	BVARPNOGO		LSTAR(7)log			VECMXPNOGO	
AR124	LSTAR(4)log	SETAR(1)log	AAR(3)log		LSTAR(2)log			BVARPDNOGO	
VARPNOGO	LSTAR(6)log	LSTAR(1)log	SETAR(9)log		AAR(8)dlog			SETAR(1)log	
LSTAR(1)log	SETAR(7)log	AAR(5)log	SETAR(1)log		SETAR(6)log			RW4	
SETAR(8)log	LSTAR(8)log	SETAR(2)log	RW		AAR(4)log			SETAR(1)dlog	
SETAR(1)log	LSTAR(3)log	SETAR(3)log	AR12		LSTAR(6)log			SETAR(4)dlog	
SETAR(4)log	LSTAR(7)log	LSTAR(1)log1	AAR(4)log		SETAR(2)log			VADongo1112	
AAR(3)log	AAR(4)log	VEPongo1112	LSTAR(1)log		LSTAR(1)log			LSTAR(4)dlog	
VECMPNOGO	SETAR(2)log	VARXPDNOGO	AAR(1)log		SETAR(3)log			BVARP	
VECMPNOGO4	SETAR(8)log	AR12	LSTAR(3)log		LSTAR(1)log0			RW	
LSTAR(4)log	SETAR(1)dlog	SETAR(1)log	SETAR(5)log		SETAR(7)log			SETAR(2)dlog	
AAR(8)log	LSTAR(5)dlog	LSTAR(6)log	SETAR(3)log		AAR(6)dlog				
LSTAR(2)log	AAR(1)dlog		SETAR(4)log		AAR(5)dlog				
AAR(2)log	VEPongo1112		SETAR(6)log		AAR(7)dlog				
LSTAR(5)log	SETAR(5)dlog		LSTAR(4)log		LSTAR(6)dlog				
LSTAR(3)log	BVARPNOGO		LSTAR(5)log		LSTAR(1)log1				
VADongo1112	AAR(6)dlog		BVARP		SETAR(6)dlog				
AAR(5)log	LSTAR(5)log		LSTAR(6)log		LSTAR(5)dlog				
AAR(4)log	AAR(8)dlog		AAR(2)log		PECM				
LSTAR(9)log	LSTAR(2)log		SETAR(7)log		SETAR(5)dlog				
SETAR(2)log	BVARPNOGO4		SETAR(2)log		BVARP				
AAR(6)log	VARongo12		SETAR(8)log		AAR(3)dlog				
BVARPNOGO	AAR(2)dlog		LSTAR(7)log		AAR(4)dlog				
VARPNOGO4	AAR(9)dlog		AR124		SETAR(1)log				
BVARPNOGO4	AAR(7)dlog		LSTAR(2)log		SETAR(1)log				
AAR(7)log	LSTAR(8)dlog		LSTAR(8)log		AAR(2)dlog				
SETAR(9)log	LSTAR(11)dlog		AAR(2)dlog		SETAR(4)dlog				
SETAR(5)dlog	LSTAR(9)dlog		BVARPNOGO4		LSTAR(4)dlog				
VEPongo1112	VARXPDNOGO4		VECMPNOGO4		LSTAR(3)dlog				
BVARP			VECMPNOGO		LSTAR(2)dlog				
BVARPD			SETAR(8)dlog		BVARPNOGO4				
BVARPDNOGO			SETAR(5)dlog		SETAR(3)dlog				
BVARPDNOGO4			VARP						
AAR(9)log			VARongo12						
PECM			VECMP						
VECMP			VARPD						
VADongo12			AAR(8)log						
RW			VARXPD						
RW4									
AAR(5)dlog									
AAR(7)dlog									
VARPD									
SETAR(4)dlog									
SETAR(6)dlog									
VECMXP									
VECMXPNOGO4									
VARXP									

Table 24: List of Models included in the Model Confidence Set for each car brand. Forecast horizon: 24 steps ahead.

D Robustness Checks: Alternative Out-Of-Sample Periods

D.1 Seasonally Adjusted data - Recession

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
BMW	VARD	VARXDNOGO4	SETAR(1)log	VADongo12	VARXNOGO4	VAR	VECongo12	VARX
	SETAR(5)log	VARDNOGO4	SETAR(8)log	VECongo12	VADongo12	VADongo12	SETAR(3)dlog	VADongo1112
	VARDNOGO4	VARD	VARXNOGO4	SETAR(1)log	VAR	VECongo1112	VARongo12	VAR
OPEL	VARDNOGO4	VADongo12	VADongo12	VECongo12	AAR(1)log	AAR(1)log	VECM	AAR(1)log
	VARXDNOGO4	VARongo12	VECongo12	VECongo1112	VECongo1112	VECM	VECMX	VARNOGO4
	LSTAR(3)dlog	VECongo12	VARongo12	VADongo12	VECM	VECMX	AAR(1)log	VARXNOGO4
VOLKSWAGEN	AAR(1)dlog	BVARD	VARDNOGO	BVARD	VARDNOGO	VARD	VARD	VARNOGO4
	VADongo1112	VARDNOGO	VARXDNOGO	VARDNOGO	VARXDNOGO	VARXD	VARXD	VARXNOGO4
	VARDNOGO	VARXDNOGO	VADongo12	VARXDNOGO	VADongo1112	VARNOGO4	VARNOGO4	VAR
CITROEN	VARDNOGO4	VECMXNOGO4	VARDNOGO4	BVARD	VARDNOGO4	VARXDNOGO4	VADongo12	VARDNOGO4
	VARXDNOGO4	VECMNOGO4	VARXDNOGO4	VARDNOGO4	VARXDNOGO4	VARDNOGO4	VARDNOGO4	VARXDNOGO4
	VARD	BVARDNOGO4	VECMXNOGO4	VARXDNOGO4	BVARD	VADongo12	VARXDNOGO4	VADongo12
FIAT	VARongo1112	AAR(8)dlog	VADongo12	VARDNOGO4	VARDNOGO4	VARD	VARD	AAR(9)dlog
	VADongo1112	VADongo1112	AAR(8)dlog	VARXDNOGO4	VARXDNOGO4	VARXD	VARXD	VARNOGO4
	VECongo1112	LSTAR(6)log	VECongo12	VARD	AAR(9)dlog	VADongo12	VARNOGO4	VARXNOGO4
TOYOTA	SETAR(2)log	AAR(1)log	AAR(9)log	AAR(9)log	AAR(9)log	VARongo1112	VARongo1112	VARongo1112
	SETAR(4)log	VECMXNOGO	AAR(1)log	AAR(1)log	VARongo1112	AAR(9)log	SETAR(1)log	AAR(9)log
	SETAR(8)log	VECMNOGO	VECMXNOGO	AAR(7)log	VARNOGO	VARNOGO	SETAR(2)log	SETAR(1)log
JAGUAR	AAR(8)log	SETAR(8)dlog	VARongo12	AAR(3)dlog	AAR(8)log	VARongo12	LSTAR(9)dlog	SETAR(3)dlog
	SETAR(1)log	LSTAR(8)dlog	VARDNOGO4	AAR(8)log	AAR(6)log	SETAR(9)dlog	SETAR(9)dlog	LSTAR(2)dlog
	SETAR(9)dlog	SETAR(9)dlog	AAR(8)log	AAR(7)log	SETAR(9)dlog	AAR(3)dlog	LSTAR(1)dlog2	SETAR(2)dlog
KIA	BVARDNOGO	VECMXNOGO4	VECMXNOGO4	BVARNOGO4	BVARDNOGO	BVARDNOGO	AAR(7)log	SETAR(1)log
	VARongo1112	VARXNOGO4	VECMNOGO4	LSTAR(8)dlog	AAR(7)log	AAR(7)log	LSTAR(8)dlog	AAR(3)dlog
	BVARDNOGO4	AAR(7)log	BVARNOGO4	AAR(2)dlog	BVARDNOGO4	SETAR(10)dlog	LSTAR(10)dlog	SETAR(4)log
MITSUBISHI	AAR(9)dlog	AAR(9)dlog	AAR(9)dlog	AAR(9)dlog	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	LSTAR(5)log	AAR(4)dlog	LSTAR(1)dlog2	VECMNOGO	VECMNOGO	AAR(7)dlog	SETAR(2)log	LSTAR(7)log
	SETAR(5)log	AAR(8)dlog	AAR(8)log	AAR(8)log	LSTAR(2)log	LSTAR(2)log	LSTAR(7)log	LSTAR(4)log
SUZUKI	VADongo12	SETAR(10)dlog	VADongo12	VECongo12	VECongo12	AAR(5)dlog	AAR(3)log	VADongo1112
	VECongo12	VECM	VECongo12	VADongo12	AAR(5)dlog	VARongo12	AAR(3)dlog	LSTAR(5)dlog
	VECM	VECMX	LSTAR(9)dlog	SETAR(10)dlog	VARDNOGO	AR12	AAR(5)dlog	VARD

Table 25: Top three models in terms of MSPE for each forecasting horizon and each car brand.

D.2 Seasonally Adjusted data - Expansion

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
BMW	AAR(6)log	BVAR	SETAR(6)dlog	AR12	AR12	AAR(7)dlog	AAR(6)dlog	AR12
	AAR(7)log	AAR(6)log	AR12	SETAR(6)dlog	SETAR(3)log	AAR(6)dlog	SETAR(3)dlog	SETAR(1)log
	BVARNOGO4	BVARNOGO4	AAR(6)log	SETAR(1)log	SETAR(1)log	AAR(5)dlog	AAR(5)dlog	SETAR(3)dlog
OPEL	LSTAR(3)dlog	VECMXNOGO	VARNOGO4	VARNOGO4	VECM	VECongo1112	VECongo1112	VECongo1112
	LSTAR(2)dlog	VECMNOGO	VARXNOGO4	VARXNOGO4	VECMX	VECMXNOGO	VECMXNOGO	VARongo1112
	LSTAR(4)dlog	SETAR(5)log	VECMXNOGO	VECongo1112	VARNOGO4	VECMNOGO	VECMNOGO	VECMXNOGO
VOLKSWAGEN	AR12	AR124	AR124	AR124	AR124	AR124	VARongo1112	AR124
	AR124	AR12	AR12	AAR(9)log	AR12	VARongo1112	AR124	VARongo1112
	AAR(1)log	VECMXNOGO	VECMXNOGO	AR12	VARongo1112	AR12	AR12	AR12
CITROEN	BVARNOGO	VAR	VARNOGO4	VARXNOGO4	VARNOGO4	VARNOGO	LSTAR(3)log	LSTAR(3)log
	AAR(2)log	VARX	VARXNOGO4	VARNOGO4	VARXNOGO4	VARXNOGO	SETAR(4)log	SETAR(4)log
	SETAR(2)log	VARNOGO	VAR	VARX	VARXNOGO	VAR	SETAR(5)log	SETAR(5)log
FIAT	AAR(1)log	AAR(3)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	AAR(3)log	AAR(1)log	AAR(3)log	SETAR(4)log	SETAR(2)log	VARongo1112	LSTAR(2)log	SETAR(1)log
	AR124	AAR(4)log	AAR(5)log	BVARNOGO	LSTAR(2)log	SETAR(2)log	VARongo1112	SETAR(2)log
TOYOTA	VARNOGO	VARNOGO	VARNOGO	VECongo1112	VECongo1112	VECongo1112	BVAR	BVAR
	VARXNOGO	VARXNOGO	VARXNOGO	VARNOGO	BVAR	BVAR	VECongo1112	BVARNOGO4
	BVARNOGO4	BVARD	VARNOGO4	VARXNOGO	BVARNOGO4	BVARNOGO4	BVARNOGO4	VECongo1112
JAGUAR	AAR(4)log	SETAR(3)dlog	AAR(3)log	AAR(3)log	SETAR(5)log	SETAR(5)log	VECongo1112	SETAR(3)log
	AAR(5)log	LSTAR(3)dlog	AAR(4)log	LSTAR(3)log	AAR(4)log	LSTAR(4)log	SETAR(5)log	SETAR(4)log
	AAR(3)log	AAR(3)log	SETAR(2)dlog	AAR(4)log	AAR(3)log	SETAR(4)log	AAR(4)log	SETAR(5)log
KIA	VARDNOGO	LSTAR(1)log	VARDNOGO4	AAR(4)log	AAR(4)log	AAR(4)log	SETAR(3)log	AAR(4)log
	LSTAR(1)log	VARDNOGO4	AAR(4)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(4)log	LSTAR(1)log
	SETAR(9)log	AR12	AAR(7)log	AAR(3)log	AAR(6)log	SETAR(3)log	AAR(2)log	AAR(3)log
MITSUBISHI	BVARDNOGO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	BVARD	LSTAR(1)log	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	BVARDNOGO4	SETAR(1)log	AAR(2)dlog	LSTAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	VARongo12
SUZUKI	RW	AAR(2)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(3)log	SETAR(2)log	SETAR(3)log
	RW4	SETAR(2)log	AAR(3)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(3)log	AAR(3)log
	BVARDNOGO4	BVARDNOGO	AAR(2)log	AAR(4)log	SETAR(1)log	SETAR(1)log	AAR(1)log	SETAR(2)log

Table 26: Top three models in terms of MSPE for each forecasting horizon and each car brand.

E Robustness Checks: Directional Accuracy

	Steps 1-6	%	Steps 7-12	%	Steps 13-24	%
SEASONALLY ADJUSTED DATA						
BMW	AR124	68.21%	AR124	66.18%	AAR(1)log	80.51%
	SETAR(3)log	66.70%	AAR(3)log	65.88%	AR124	80.28%
	SETAR(8)log	66.20%	LSTAR(1)log	65.38%	VARongo1112	77.58%
OPEL	VADongo12	68.95%	VADongo12	69.02%	SETAR(11)dlog	77.55%
	SETAR(11)dlog	66.51%	SETAR(11)dlog	68.13%	LSTAR(8)dlog	75.06%
	VECM	65.15%	LSTAR(8)dlog	67.88%	BVARNOGO	74.12%
VOLKSWAGEN	AR124	74.53%	AR124	82.88%	VARongo1112	84.59%
	VARongo1112	73.79%	VARongo1112	80.79%	AR124	84.25%
	AR12	72.28%	AR12	80.66%	AR12	82.25%
CITROEN	VECMXNOGO4	65.59%	VECongo1112	67.69%	VECongo1112	75.12%
	VECMNOGO4	65.59%	VAR	67.31%	AAR(4)dlog	66.61%
	VECM	65.15%	VARX	67.31%	AAR(5)dlog	61.16%
FIAT	VECongo1112	61.02%	VECongo1112	74.54%	AAR(1)dlog	85.16%
	VADongo12	60.41%	VECongo12	67.31%	AAR(2)dlog	83.39%
	VARongo1112	59.99%	LSTAR(12)dlog	64.59%	AAR(3)dlog	83.21%
TOYOTA	VECongo1112	66.56%	VECongo12	71.59%	BVAR	74.53%
	VARongo1112	65.27%	VADongo12	71.57%	BVARNOGO4	72.44%
	VECMXNOGO	64.97%	BVARNOGO4	71.06%	VADongo12	71.03%
JAGUAR	LSTAR(3)log	68.15%	SETAR(4)log	66.42%	VARongo1112	70.79%
	AAR(4)log	68.13%	LSTAR(3)log	66.17%	AR12	70.10%
	LSTAR(4)log	68.12%	LSTAR(4)log	65.63%	VARongo12	69.98%
KIA	AAR(7)log	67.52%	AAR(2)log	71.68%	SETAR(5)log	78.77%
	AAR(4)log	66.06%	AAR(3)log	71.40%	SETAR(6)log	78.77%
	VARXNOGO4	65.80%	AAR(6)log	71.11%	AAR(2)log	78.46%
MITSUBISHI	VECongo1112	69.07%	VECongo12	81.34%	VECongo1112	84.41%
	VECongo12	67.31%	VARongo12	77.16%	VECongo12	82.43%
	AAR(4)dlog	65.86%	VECongo1112	77.16%	VARongo12	79.91%
SUZUKI	AAR(1)log	71.10%	SETAR(3)log	75.20%	AAR(1)log	84.56%
	SETAR(3)log	68.56%	SETAR(2)log	74.67%	SETAR(2)log	79.60%
	SETAR(2)log	68.08%	AAR(1)log	73.82%	SETAR(3)log	77.50%
RAW DATA						
	Steps 1-6	**%**	**Steps 7-12**	**%**	**Steps 13-24**	**%**
BMW	VADongo1112	84.45%	VADongo1112	82.27%	LSTAR(10)log	81.52%
	VARongo12	84.38%	BVARPD	79.96%	LSTAR(11)log	80.57%
	BVARPD	83.89%	LSTAR(11)log	79.80%	LSTAR(9)log	80.44%
OPEL	VADongo1112	87.02%	VEPongo1112	81.69%	VARongo1112	81.99%
	VARongo1112	86.41%	VADongo1112	81.13%	VADongo1112	77.82%
	VADongo12	85.48%	VADongo12	79.77%	VEPongo1112	77.76%
VOLKSWAGEN	BVARPNOGO	82.85%	VEPongo1112	78.75%	VECPongo12	84.64%
	VECMP	82.61%	BVARPNOGO	78.42%	VEPongo1112	81.13%
	VECMXP	82.61%	VECMP	76.48%	VARongo1112	77.99%
CITROEN	VARPNOGO4	80.54%	VEPongo1112	76.72%	AAR(3)log	78.02%
	VARXPNOGO4	80.54%	BVARPNOGO4	76.52%	AAR(4)log	77.40%
	VARPNOGO	80.29%	BVARPNOGO	76.49%	LSTAR(3)dlog	76.97%
FIAT	VARongo1112	85.77%	PECM	84.58%	VARongo1112	83.02%
	VADongo1112	84.65%	VADongo12	79.70%	VEPongo1112	80.62%
	BVARP	84.37%	VADongo1112	78.25%	VADongo12	79.48%
TOYOTA	VARongo1112	81.73%	BVARP	79.23%	VEPongo1112	79.42%
	VEPongo1112	79.64%	VADongo12	78.56%	VARongo1112	79.03%
	BVARPNOGO4	77.02%	VECPongo12	78.52%	BVARP	75.00%
JAGUAR	VARXPNOGO4	80.60%	VARXPNOGO4	79.57%	VARPNOGO	80.45%
	VARXPNOGO	80.05%	VARPNOGO	79.12%	VARPNOGO4	77.87%
	AR12	80.04%	VARPNOGO4	77.58%	AR12	77.32%
KIA	BVARPDNOGO	72.52%	VARPNOGO4	67.93%	VARPNOGO4	70.62%
	VARongo1112	71.97%	VARXPNOGO4	67.93%	VARXPNOGO4	70.62%
	VADongo1112	71.96%	VEPongo1112	66.56%	RW	69.21%
MITSUBISHI	VEPongo1112	75.45%	VEPongo1112	77.87%	VEPongo1112	82.60%
	VARongo1112	74.56%	VECPongo12	76.31%	VECPongo12	82.07%
	VECPongo12	72.98%	LSTAR(2)log	73.21%	VARongo1112	80.77%
SUZUKI	VARongo1112	75.88%	SETAR(6)log	77.95%	AAR(2)log	78.95%
	BVARPNOGO	73.27%	SETAR(5)log	77.55%	SETAR(8)log	78.90%
	AR12	72.68%	AAR(4)log	76.77%	VARongo1112	78.75%

Table 27: Top three models in terms of average directional accuracy (in %) for each car brand, for short-term forecasts (1-6 steps ahead), medium-term forecasts (7-12 steps ahead), and long-term forecasts (13-24 steps ahead).

F Robustness Checks: Variability of Google Data

SEASONALLY ADJUSTED DATA										
	BMW	Opel	Volkswagen	Citroen	Fiat	Toyota	Jaguar	Kia	Mitsubishi	Suzuki
VECM	1.08	3.65	0.94	1.00	0.65	NC	1.03	0.95	1.06	1.95
VECMX	1.07	3.65	0.94	1.00	0.65	NC	0.85	NC	0.77	1.95
VAR	0.92	0.66	1.00	1.00	1.15	0.96	1.00	0.98	0.95	1.18
VARD	1.14	1.61	0.79	0.22	1.49	1.16	0.94	0.92	0.99	1.26
VARX	0.97	0.66	1.00	1.00	1.07	0.86	1.42	0.97	0.95	1.18
VARXD	1.07	1.61	0.79	1.00	1.41	1.05	1.00	1.28	1.00	1.26
VECongo1112	1.02	0.94	1.03	0.99	0.99	1.13	1.02	1.05	0.95	1.05
VECongo12	1.00	0.99	1.04	1.05	1.48	0.41	0.87	1.14	1.01	1.10
VARongo1112	1.00	1.02	0.98	0.99	1.03	1.04	0.98	1.04	1.03	0.89
VARongo12	1.01	0.97	1.04	1.09	1.01	1.19	0.83	1.05	1.03	0.91
VADongo1112	1.02	0.97	1.00	1.00	0.99	1.03	0.96	2.15	1.00	0.98
VADongo12	1.08	0.95	1.04	0.95	0.97	1.02	1.13	3.58	0.93	1.19
BVAR	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.97	1.00	1.00
BVARD	1.00	0.99	1.00	1.00	1.00	1.00	1.00	2.06	0.99	0.99
RAW DATA										
	BMW	Opel	Volkswagen	Citroen	Fiat	Toyota	Jaguar	Kia	Mitsubishi	Suzuki
VECMP	0.61	0.99	1.54	1.20	0.25	NC	1.03	NC	1.57	0.58
VECMXP	1.16	0.99	1.54	1.20	0.25	NC	0.81	NC	1.10	0.58
VEPongo1112	1.00	0.99	0.97	0.98	0.97	0.83	1.04	1.00	0.96	1.04
VECPongo12	0.96	0.93	0.92	1.59	1.27	0.90	1.04	1.07	1.06	2.73
VARP	0.96	0.97	0.98	1.07	0.96	0.99	0.97	1.00	1.08	0.95
VARPD	1.06	1.00	0.93	1.01	1.17	0.57	1.00	0.80	1.12	1.22
VARXP	1.12	0.97	0.98	1.07	0.96	0.88	0.98	1.00	0.97	0.95
VARXPD	1.09	1.00	0.93	1.01	1.17	1.26	1.00	0.80	1.25	1.22
VADongo1112	1.02	1.00	0.96	1.00	0.96	1.49	1.00	1.00	1.00	1.05
VADongo12	1.04	0.91	1.04	1.21	0.98	1.10	0.98	1.03	1.00	3.69
VARongo1112	0.94	0.96	0.96	0.99	0.96	0.88	1.00	0.98	0.99	1.04
VARongo12	1.02	0.93	1.08	1.47	1.00	1.25	0.97	1.02	0.97	1.20
BVARP	1.00	1.00	1.00	1.00	0.99	0.95	1.00	1.00	1.00	1.00
BVARPD	1.00	1.00	1.00	1.00	0.98	0.99	1.00	1.00	1.00	1.00
PECM	0.97	1.46	1.70	1.13	1.00	3.10	1.04	0.63	1.10	1.25

Table 28: Average ratio – averaged across all forecasting horizons – of the MSE for the forecasts computed with GIs downloaded on the 15/08/2014, with respect to MSE for the forecasts computed with the average GIs downloaded between 15/08/2014 and 02/09/2014. NC=not converged.

G Robustness Checks: Additional Car Brands

G.1 Seasonally Adjusted data

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12
AUDI	BVARNOGO	AAR(3)log	BVARNOGO	VARXNOGO	LSTAR(6)log	AAR(5)log	AAR(5)log	AAR(5)log	AAR(5)log	SETAR(3)log	SETAR(3)log	SETAR(2)log
	BVARNOGO4	BVARNOGO	AAR(3)log	BVARNOGO	LSTAR(7)log	LSTAR(7)log	SETAR(1)log	SETAR(6)log	SETAR(1)log	SETAR(2)log	SETAR(2)log	AAR(7)log
	BVAR	VARXNOGO	BVAR	BVAR	SETAR(2)log	SETAR(2)log	SETAR(6)log	SETAR(5)log	LSTAR(4)log	AAR(5)log	AAR(6)log	LSTAR(4)log
FORD	BVARDNOGO	BVARDNOGO	BVARDNOGO	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	VARongo1112	VARongo1112
	BVARDNOGO4	BVARDNOGO4	BVARD	AAR(3)log	AAR(5)log	AAR(5)log	AAR(5)log	VARongo1112	AAR(5)log	VARongo1112	AAR(2)log	AAR(2)log
	BVARD	BVARD	BVARDNOGO4	AAR(5)log	AAR(4)log	AAR(4)log	VARongo1112	AAR(5)log	VARongo1112	AAR(5)log	AAR(5)log	AAR(5)log
HYUNDAI	VARXDNOGO	VARDNOGO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo12	VADongo12	VECongo1112
	LSTAR(1)dlog	VARXDNOGO	VARDNOGO	VARDNOGO	AR12	AR12	AR12	AR12	VADongo12	VADongo12	VECongo12	VADongo12
	VARDNOGO	VARDNOGO4	LSTAR(1)dlog	AAR(2)log	VADongo1112	VADongo1112	AAR(1)log	VARDNOGO	AR12	VECongo1112	VECongo1112	VECongo12
MAZDA	SETAR(1)log	AAR(6)dlog	AAR(6)dlog	SETAR(6)dlog	VECongo12	VECongo12	VECongo12	VECongo12	VECongo12	VECongo12	VECongo12	VECongo12
	SETAR(2)log	AAR(5)dlog	SETAR(6)dlog	LSTAR(6)dlog	VADongo12	VADongo12	VADongo12	VADongo12	VADongo12	VADongo12	VADongo12	VADongo12
	RW	VARongo1112	LSTAR(6)dlog	AAR(6)dlog	SETAR(5)dlog	SETAR(5)dlog	SETAR(6)dlog	SETAR(6)dlog	SETAR(6)dlog	LSTAR(1)dlog2	LSTAR(1)dlog2	LSTAR(1)dlog2
NISSAN	LSTAR(1)log	VADongo1112	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	AAR(2)log	AAR(3)log	AAR(3)log	LSTAR(1)log	LSTAR(1)log	AAR(1)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log
	AAR(3)log	AAR(2)log	LSTAR(1)log	AAR(3)log	AAR(3)log	LSTAR(1)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log
PEUGEOUT	BVARD	BVARD	BVARD	BVARNOGO4	BVARNOGO4	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	BVARDNOGO4	BVARDNOGO	BVARDNOGO	BVAR	BVAR	BVAR	AAR(3)log	AAR(3)log	VADongo1112	AAR(5)dlog	AAR(5)dlog	SETAR(3)dlog
	VADongo1112	BVARDNOGO4	BVARDNOGO4	BVARNOGO	BVARNOGO	BVARNOGO4	AAR(2)log	VADongo1112	SETAR(3)dlog	VADongo1112	SETAR(3)dlog	LSTAR(8)dlog
RENAULT	LSTAR(1)log	LSTAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	VECMXNOGO4	VARXNOGO	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	SETAR(1)log	SETAR(2)log	VARNOGO	VARNOGO	VARXNOGO	VARXNOGO	VECMXNOGO4	VARXNOGO	SETAR(8)log	SETAR(8)log	LSTAR(6)log	LSTAR(6)log
	SETAR(2)log	VARDNOGO	VARDNOGO	VARXNOGO	VARNOGO	AAR(1)log	AAR(1)log	BVARNOGO	LSTAR(6)log	LSTAR(6)log	SETAR(7)log	SETAR(7)log
HONDA	BVARNOGO	VECMXNOGO	VECongo1112	BVARNOGO	BVARNOGO	BVAR	BVAR	SETAR(9)log	SETAR(9)log	SETAR(9)log	VECongo1112	VECongo1112
	BVARNOGO4	VECongo1112	BVARNOGO	BVARNOGO4	BVARNOGO4	BVARNOGO4	BVARNOGO	VECMXNOGO	SETAR(8)log	VECMXNOGO	VECMXNOGO	VECMXNOGO
	VARXNOGO	BVARNOGO	VARongo1112	VECongo1112	BVAR	BVARNOGO	BVARNOGO4	SETAR(8)log	VECMXNOGO	VECongo1112	SETAR(9)log	SETAR(9)log
LAND ROVER	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	VARDNOGO	AAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog	AAR(2)dlog	AAR(2)dlog	AAR(1)dlog
	VARDNOGO4	VARDNOGO4	AAR(1)dlog	AAR(1)dlog	SETAR(3)dlog	AAR(3)dlog	SETAR(3)dlog	AAR(1)dlog	LSTAR(1)dlog	SETAR(1)dlog	AAR(1)dlog	SETAR(6)dlog
	VARXDNOGO	VARD	SETAR(3)dlog	SETAR(3)dlog	AAR(1)dlog	AAR(2)dlog	VARDNOGO	VARDNOGO	AAR(1)dlog	AAR(1)dlog	AAR(3)dlog	SETAR(3)dlog
PORSCHE	LSTAR(1)dlog	AAR(4)dlog	LSTAR(1)dlog	SETAR(6)log	SETAR(6)log	SETAR(6)log	SETAR(6)log	SETAR(6)log	AAR(5)dlog	SETAR(6)log	SETAR(6)log	SETAR(7)dlog
	AAR(4)dlog	AAR(5)dlog	SETAR(2)dlog	SETAR(7)log	LSTAR(3)dlog	SETAR(7)log	SETAR(7)log	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog
	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	LSTAR(3)dlog	SETAR(7)log	AAR(4)dlog	AAR(4)dlog	AAR(6)dlog	SETAR(6)log	SETAR(7)dlog	SETAR(7)dlog	SETAR(9)dlog
SUBARU	LSTAR(1)log	LSTAR(1)log	SETAR(1)log	VARNOGO4	VARNOGO4	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	SETAR(1)log	SETAR(1)log	LSTAR(1)log	AAR(1)log	AAR(1)log	SETAR(1)log	VARNOGO4	VARNOGO4	VARNOGO4	SETAR(9)log	SETAR(2)log	SETAR(2)log
	AAR(1)log	AAR(3)log	AAR(1)log	LSTAR(1)log	SETAR(1)log	VARNOGO4	SETAR(2)log	SETAR(2)log	SETAR(9)log	VARNOGO	SETAR(7)log	SETAR(7)log
VOLVO	AAR(2)dlog	VECongo1112	SETAR(4)log	VECongo1112	VECongo1112	VECongo1112	VECM	VECM	SETAR(3)log	VECongo1112	VECongo1112	VECongo1112
	AAR(3)dlog	AAR(2)dlog	AAR(2)dlog	VECM	SETAR(6)dlog	SETAR(3)log	VECongo1112	VECongo1112	VECongo1112	VECM	SETAR(3)log	VARXDNOGO
	AR12	SETAR(2)dlog	SETAR(2)dlog	SETAR(2)dlog	SETAR(4)dlog	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)log	VECM	VECMNOGO4	VECM	VARongo1112

	Step 13	Step 14	Step 15	Step 16	Step 17	Step 18	Step 19	Step 20	Step 21	Step 22	Step 23	Step 24
AUDI	LSTAR(4)log	LSTAR(4)log	AAR(5)log	AAR(6)log	LSTAR(4)log	LSTAR(4)log	LSTAR(4)log	LSTAR(4)log	SETAR(5)log	LSTAR(3)log	LSTAR(3)log	LSTAR(3)log
	LSTAR(3)log	LSTAR(3)log	SETAR(2)log	AAR(7)log	LSTAR(6)log	SETAR(1)log	LSTAR(3)log	LSTAR(3)log	SETAR(2)log	LSTAR(4)log	LSTAR(4)log	LSTAR(4)log
	AAR(3)log	AAR(3)log	LSTAR(4)log	SETAR(2)log	LSTAR(7)log	LSTAR(3)log	SETAR(1)log	SETAR(5)log	LSTAR(4)log	AAR(3)log	LSTAR(2)log	LSTAR(2)log
FORD	AAR(5)log	AAR(4)log	AAR(4)log	AAR(4)log	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log	AR124	AR124	AR124	AR124
	VARongo1112	AAR(3)log	AAR(3)log	AAR(3)log	AAR(4)log	AAR(4)log	AR124	AR124	AAR(3)log	AAR(3)log	AAR(3)log	AAR(3)log
	AAR(3)log	AAR(5)log	AAR(5)log	AAR(5)log	AAR(5)log	AR124	AAR(4)log	AAR(4)log	AAR(6)log	AAR(4)log	AAR(6)log	AAR(4)log
HYUNDAI	VADongo12	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	VECongo1112	VADongo12	AR12	VADongo12	AR12	VADongo12	AR12	AR12	VADongo12	AR12	AR12	AR12
	AR12	AR12	VADongo12	AR12	VADongo1112	AR12	VADongo12	VADongo12	AR12	LSTAR(2)dlog	LSTAR(1)dlog	VARXDNOGO
MAZDA	VECongo12	SETAR(10)dlog	SETAR(6)dlog	AAR(6)dlog	LSTAR(1)dlog2	VARongo12	VARongo12	LSTAR(1)dlog2	LSTAR(1)dlog2	VECongo12	VECongo12	VECongo12
	AAR(6)dlog	AAR(6)dlog	SETAR(7)dlog	AAR(5)dlog	AAR(6)dlog	LSTAR(1)dlog2	LSTAR(1)dlog2	VECongo12	VECongo12	LSTAR(1)dlog2	LSTAR(1)dlog2	LSTAR(1)dlog2
	VADongo12	AAR(4)dlog	AAR(6)dlog	AAR(4)dlog	AAR(5)dlog	AAR(6)dlog	VECongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12
NISSAN	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log
	AAR(2)log	AAR(2)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log
	AAR(3)log	AAR(3)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log
PEUGEOUT	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VADongo12	VECongo1112	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog
	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog	VADongo12	VADongo12	VECongo1112	SETAR(3)dlog	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112
	AAR(5)dlog	SETAR(2)dlog	VADongo12	SETAR(3)dlog	SETAR(3)dlog	SETAR(3)dlog	VADongo12	SETAR(5)dlog	SETAR(5)dlog	SETAR(5)dlog	SETAR(2)dlog	SETAR(5)dlog
RENAULT	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log	LSTAR(6)log
	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log	SETAR(7)log
	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log	SETAR(8)log
HONDA	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VARongo12	VARongo12	VARongo12	VARongo12	SETAR(1)log
	SETAR(9)log	VECMXNOGO	SETAR(9)log	SETAR(9)log	SETAR(9)log	SETAR(9)log	VARongo12	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VARongo12
	VARongo12	SETAR(9)log	VECongo12	SETAR(7)log	SETAR(7)log	VECMXNOGO	SETAR(9)log	VECongo12	VECongo12	VECongo12	VECongo12	VECongo1112
LAND ROVER	AAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog	SETAR(1)dlog
	SETAR(1)dlog	AAR(1)dlog	AAR(1)dlog	AAR(7)dlog	AAR(7)dlog	AAR(7)dlog	AAR(7)dlog	AAR(7)dlog	AAR(6)dlog	AAR(7)dlog	AAR(6)dlog	AAR(1)dlog
	SETAR(3)dlog	RW	AAR(7)dlog	AAR(1)dlog	AAR(1)dlog	AAR(1)dlog	AAR(1)dlog	AAR(1)dlog	AAR(7)dlog	SETAR(9)dlog	AAR(1)dlog	AAR(7)dlog
PORSCHE	SETAR(9)dlog	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog	AAR(4)dlog	AAR(3)log	SETAR(7)dlog	SETAR(7)dlog	AAR(8)dlog	AAR(8)dlog
	SETAR(7)dlog	SETAR(9)dlog	SETAR(6)dlog	VARX	AAR(2)dlog	AAR(6)dlog	AAR(8)dlog	AAR(8)dlog	AAR(4)log	AAR(8)dlog	SETAR(7)dlog	SETAR(7)dlog
	AAR(4)dlog	SETAR(7)dlog	SETAR(7)dlog	AAR(5)dlog	AAR(6)dlog	AAR(5)dlog	AAR(6)dlog	AAR(4)log	AAR(3)log	AAR(6)dlog	AAR(6)dlog	AAR(4)dlog
SUBARU	AAR(1)log	AAR(1)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	AAR(1)log	AAR(1)log
	SETAR(2)log	SETAR(2)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	SETAR(1)log	AAR(1)log	AAR(1)log	SETAR(1)log	SETAR(4)log
	AR12	SETAR(9)log	SETAR(3)log	LSTAR(2)log	SETAR(4)log	SETAR(3)log	LSTAR(1)log	AAR(1)log	SETAR(3)log	SETAR(4)log	SETAR(2)log	SETAR(3)log
VOLVO	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	VECongo1112	AR12
	SETAR(3)log	LSTAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	SETAR(1)log	AR12	AAR(1)log
	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	AAR(1)log	AAR(1)log	AAR(1)log	VECongo1112

Table 29: Top three models in terms of MSPE for each forecasting horizon and each car brand.

G.2 Raw data

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12
AUDI	VARongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARPNOGO	VARPNOGO	AR124	VEPongo1112
	VEPongo1112	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VARPNOGO	VEPongo1112	VEPongo1112	AR12	AR12
	VARPNOGO	VARongo1112	VARP	VARXPNOGO	VARXPNOGO	VADongo1112	VARXPNOGO	VADongo1112	AR12	AR12	VEPongo1112	AR124
FORD	VEPongo1112	VARongo1112	VEPongo1112	VARongo1112	VARongo1112	VARongo1112	AAR(5)log	AAR(5)log	SETAR(1)log	LSTAR(1)log1	AAR(6)log	AAR(6)log
	VADongo1112	VEPongo1112	VARongo1112	VARongo12	LSTAR(2)log	AAR(5)log	VARongo1112	SETAR(2)log	SETAR(2)log	SETAR(4)log	LSTAR(2)log	LSTAR(9)log
	VARongo1112	VARPNOGO	VARongo12	VEPongo1112	AAR(5)log	SETAR(1)log	SETAR(4)log	AAR(4)log	LSTAR(2)log	LSTAR(2)log	AAR(5)log	AAR(4)log
HYUNDAI	VEPongo1112	VEPongo1112	SETAR(11)dlog	AR124	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARongo12	VADongo1112	VADongo1112
	AR12	AR12	AR12	VEPongo1112	VARongo1112	SETAR(11)dlog	SETAR(11)dlog	VADongo1112	VADongo1112	VADongo12	VEPongo1112	RW
	VADongo1112	VADongo1112	VEPongo1112	VARongo1112	VADongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VEPongo1112	VARongo12	VEPongo1112
MAZDA	VARPNOGO	VARongo1112	VARongo1112	VARongo1112	VECPongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo1112	VARongo1112	VARongo1112
	VARXPNOGO	VARPNOGO	VARPNOGO	VARongo12	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo12	VARongo12	VARongo12
	VARongo1112	VARXPNOGO	VARXPNOGO	VECPongo12	VARongo12	VECPongo12	VADongo12	VADongo12	VADongo1112	VADongo1112	VADongo1112	VADongo12
NISSAN	AR12	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AAR(2)log	AR12	AR12	AR12	AR12	VADongo1112
	VARP	AAR(3)log	AAR(3)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(1)log	AAR(2)log	AAR(2)log	AAR(2)log	BVARP	VEPongo1112
	VARPNOGO	AR12	AAR(1)log	AAR(3)log	AAR(3)log	AR12	AR12	AAR(1)log	AAR(1)log	AAR(1)log	VEPongo1112	BVARP
PEUGEOUT	VADongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VADongo1112	VADongo1112	VADongo1112	SETAR(4)dlog	SETAR(4)dlog	SETAR(10)dlog	SETAR(10)dlog	SETAR(10)dlog
	VEPongo1112	VADongo1112	VADongo1112	VADongo1112	VEPongo1112	VEPongo1112	VEPongo1112	SETAR(10)dlog	SETAR(10)dlog	SETAR(4)dlog	SETAR(4)dlog	AAR(1)dlog
	VARPDNOGO4	VARongo1112	VECMXPNOGO	VECMXPNOGO	SETAR(10)dlog	SETAR(10)dlog	SETAR(10)dlog	SETAR(9)dlog	SETAR(5)dlog	SETAR(9)dlog	AAR(1)dlog	SETAR(4)dlog
RENAULT	VEPongo1112	VADongo12	VADongo12	SETAR(2)log	SETAR(3)log	SETAR(1)log	LSTAR(1)log	SETAR(3)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	AAR(4)log
	VARPNOGO	VECPongo12	SETAR(1)log	SETAR(3)log	SETAR(1)log	LSTAR(1)log	SETAR(1)log	AAR(4)log	AAR(4)log	SETAR(3)log	SETAR(5)log	LSTAR(1)log
	VARXPNOGO	VECMXPNOGO	VECMXPNOGO	AAR(2)log	LSTAR(1)log	SETAR(3)log	SETAR(3)log	LSTAR(1)log	SETAR(1)log	SETAR(5)log	SETAR(3)log	SETAR(5)log
HONDA	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VARongo1112	VARongo12	VARongo12
	VEPongo1112	VECMXPNOGO	VEPongo1112	VECMXPNOGO	VECMXPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VARongo12	VARongo1112	VARongo1112
	VARPNOGO	VECMPNOGO	VECMXPNOGO	VECMPNOGO	VECMPNOGO	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VECMXPNOGO	VECPongo12	VECPongo12
LAND ROVER	AAR(1)dlog	AAR(1)dlog	AAR(1)dlog	AAR(3)dlog	VEPongo1112	VEPongo1112	PECM	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	RW
	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	PECM	AAR(2)dlog	VARPNOGO	VARPNOGO	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	AAR(2)dlog
	LSTAR(1)dlog	RW	AAR(3)dlog	RW	AAR(2)dlog	PECM	VEPongo1112	AAR(3)dlog	VARPNOGO	VADongo12	VADongo12	AAR(3)dlog
PORSCHE	VARongo1112	VARongo1112	LSTAR(11)dlog	LSTAR(11)dlog	VARongo1112	AR12	AR12	AR12	AR12	AR12	LSTAR(11)dlog	VADongo1112
	LSTAR(11)dlog	VEPongo1112	VARongo1112	VARongo1112	AR12	AR124	AR124	LSTAR(11)dlog	LSTAR(11)dlog	LSTAR(11)dlog	AR12	AR12
	AR12	BVARPNOGO	SETAR(11)dlog	SETAR(11)dlog	AR124	VARongo1112	VARongo1112	AR124	AR124	AR124	SETAR(11)dlog	AR124
SUBARU	VADongo1112	VARP	SETAR(1)log	LSTAR(1)log	SETAR(4)log	SETAR(4)log	SETAR(4)log	SETAR(4)log	SETAR(1)log	SETAR(1)log	AAR(9)log	BVARPDNOGO
	VEPongo1112	VARXP	LSTAR(1)log	SETAR(1)log	LSTAR(1)log	SETAR(3)log	SETAR(8)log	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	SETAR(4)log	AAR(9)log
	VARongo1112	LSTAR(1)log	AAR(3)log	SETAR(3)log	SETAR(1)log	LSTAR(1)log	LSTAR(1)log	SETAR(8)log	SETAR(4)log	SETAR(8)log	LSTAR(1)log	AAR(8)log
VOLVO	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	AAR(7)log
	VARongo1112	VADongo1112	PECM	PECM	VECMPNOGO	VECMPNOGO	PECM	VECMPNOGO4	VECMP	VECMP	AAR(7)log	BVARPNOGO4
	PECM	VARongo1112	VADongo1112	VARongo1112	PECM	PECM	VECMPNOGO	VECMPNOGO	AAR(6)log	AAR(6)log	BVARPNOGO4	AAR(6)log

	Step 13	Step 14	Step 15	Step 16	Step 17	Step 18	Step 19	Step 20	Step 21	Step 22	Step 23	Step 24
AUDI	VEPongo1112	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	BVARPNOGO	BVARPNOGO
	VECMPNOGO	VECMPNOGO4	VEPongo1112	VECMP	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VECMPNOGO	VECMPNOGO
	VECMPNOGO4	VEPongo1112	VECMPNOGO4	VEPongo1112	VARPNOGO	VARPNOGO	VARPNOGO	BVARPNOGO	BVARPNOGO	BVARPNOGO	VEPongo1112	VEPongo1112
FORD	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	AAR(6)log	AAR(6)log	LSTAR(1)log1	LSTAR(1)log1	VARongo12	VEPongo1112	VEPongo1112	LSTAR(1)log1	SETAR(3)log	SETAR(2)log	VARongo12	SETAR(1)log
	AAR(5)log	AAR(5)log	AAR(6)log	VEPongo1112	LSTAR(3)log	SETAR(2)log	LSTAR(1)log1	VEPongo1112	LSTAR(9)log	SETAR(3)log	SETAR(2)log	SETAR(3)log
HYUNDAI	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	VEPongo1112	SETAR(11)dlog	SETAR(11)dlog	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VEPongo1112	VARongo12	VECMXPNOGO	VECMXPNOGO
	VADongo1112	SETAR(10)dlog	VEPongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VADongo1112	VEPongo1112	VEPongo1112	VEPongo1112
MAZDA	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112
	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12
	VADongo12	VADongo12	VADongo1112	VECPongo12	VECPongo12	VECPongo12	VADongo12	VEPongo1112	VEPongo1112	VECPongo12	VECPongo12	VECPongo12
NISSAN	BVARP	BVARP	BVARP	LSTAR(1)log	AAR(2)log	AAR(2)log	AAR(2)log	VARPDNOGO	AAR(2)log	PECM	VARPDNOGO	VARPDNOGO4
	VEPongo1112	AAR(2)log	AAR(2)log	AAR(2)log	VARPDNOGO	AAR(1)log	AAR(1)log	AAR(2)log	AAR(1)log	AAR(2)log	AAR(2)log	VARPDNOGO
	VADongo1112	VEPongo1112	AAR(1)log	AAR(1)log	AAR(1)log	VARongo1112	VARPDNOGO	AAR(1)log	PECM	AAR(1)log	AAR(1)log	PECM
PEUGEOUT	SETAR(10)dlog	SETAR(10)dlog	SETAR(10)dlog	SETAR(10)dlog	SETAR(10)dlog	SETAR(4)dlog	SETAR(4)dlog	SETAR(4)dlog	SETAR(4)dlog	SETAR(4)dlog	SETAR(4)dlog	SETAR(4)dlog
	SETAR(4)dlog	AR124	SETAR(11)dlog	SETAR(4)dlog	SETAR(9)dlog	SETAR(9)dlog	SETAR(9)dlog	SETAR(9)dlog	SETAR(5)dlog	SETAR(5)dlog	SETAR(5)dlog	SETAR(9)dlog
	AAR(1)dlog	SETAR(11)dlog	SETAR(4)dlog	SETAR(9)dlog	SETAR(4)dlog	SETAR(10)dlog	SETAR(5)dlog	SETAR(5)dlog	SETAR(9)dlog	SETAR(9)dlog	SETAR(9)dlog	SETAR(7)dlog
RENAULT	SETAR(5)log	LSTAR(1)log	SETAR(6)log	AAR(4)log	SETAR(5)log	SETAR(4)log	SETAR(4)log	SETAR(5)log	SETAR(6)log	SETAR(9)log	SETAR(8)log	SETAR(5)log
	LSTAR(1)log	SETAR(3)log	AAR(4)log	LSTAR(1)log	SETAR(4)log	SETAR(5)log	SETAR(5)log	SETAR(4)log	SETAR(5)log	SETAR(5)log	SETAR(9)log	SETAR(4)log
	SETAR(4)log	AAR(4)log	LSTAR(1)log	SETAR(4)log	AAR(4)log	AAR(4)log	AAR(4)log	SETAR(7)log	SETAR(4)log	SETAR(4)log	SETAR(6)log	LSTAR(1)log
HONDA	VARongo12	VARongo1112	VARongo12	VARongo12	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo1112	VARongo12	VARongo12
	VARongo1112	VARongo12	VARongo1112	VARongo1112	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo12	VARongo1112	VARongo1112
	VECPongo12	VECMXPNOGO	VECPongo12	VECPongo12	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECPongo12
LAND ROVER	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog	AAR(2)dlog
	AAR(3)dlog	RW	RW	RW	RW	RW	RW	RW	RW	RW	RW	RW
	VADongo1112	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	VARXPNOGO4	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	AAR(3)dlog	AAR(1)dlog
PORSCHE	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VADongo12	VADongo12	VADongo12	VECMXPNOGO	AR124	VARPNOGO	VARPNOGO
	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMPNOGO	VECMXPNOGO	VECMXPNOGO	VECMXPNOGO	VECMPNOGO	AR12	VARXPNOGO	VARXPNOGO
	PECM	VADongo12	VADongo12	BVARPNOGO	VADongo12	VECMPNOGO	VECMPNOGO	VECMPNOGO	AR124	VADongo12	VADongo12	VADongo12
SUBARU	SETAR(8)log	AAR(5)log	SETAR(8)log	AAR(8)log	AAR(8)log	AR12	SETAR(1)log	AAR(8)log	AR12	SETAR(1)log	SETAR(2)log	SETAR(3)log
	LSTAR(1)log	AR12	AAR(5)log	AAR(5)log	LSTAR(11)dlog	AAR(4)log	AR12	AR12	AR124	LSTAR(1)log	SETAR(1)log	SETAR(1)log
	SETAR(1)log	SETAR(8)log	SETAR(7)log	AAR(6)log	AR12	AAR(6)log	SETAR(2)log	SETAR(1)log	SETAR(2)log	SETAR(4)log	SETAR(4)log	LSTAR(1)log
VOLVO	BVARPNOGO4	VEPongo1112	LSTAR(1)log	LSTAR(1)log	LSTAR(1)log	SETAR(1)log	LSTAR(1)log	LSTAR(1)log	SETAR(1)log	LSTAR(1)log	SETAR(2)log	SETAR(2)log
	VEPongo1112	BVARPNOGO4	LSTAR(2)log	LSTAR(2)log	SETAR(1)log	LSTAR(1)log	SETAR(1)log	SETAR(1)log	LSTAR(1)log	SETAR(1)log	SETAR(1)log	AAR(2)log
	AAR(1)log	LSTAR(2)log	SETAR(1)log	SETAR(1)log	LSTAR(2)log	LSTAR(2)log	AAR(2)log	SETAR(2)log	SETAR(2)log	SETAR(2)log	LSTAR(1)log	AAR(1)log

Table 30: Top three models in terms of MSPE for each forecasting horizon and each car brand.

Forecasting German Car Sales Using Google Data and Multivariate Models

Technical appendix - not for publication. To be posted on the authors' web sites.

Dean Fantazzini   Zhamal Toktamysova

A Unit Root Tests allowing for Structural Breaks

A.1 Minimum LM Unit Root test by Lee and Strazicich (2003)

Unit root tests where the break dates are data-driven are called unit root tests with endogenous structural breaks. Zivot and Andrews (1992) consider the case of one break, while Lumsdaine and Papell (1997) of two breaks. Both tests assume no break under the null of a unit root. However, if the data generating process has a unit root with break(s), then these tests exhibit size distortions. In this regard, Nunes, Newbold, and Kuan (1997) and Lee and Strazicich (2001) point out that this frequently leads to a spurious rejection of the null hypothesis, so that the case of stationarity with break(s) is accepted too often. Moreover, the break point is often incorrectly determined one period prior to the true break date (Lee and Strazicich (2001)). The one-break and two-break Langrange Multiplier (LM) unit root tests by Lee and Strazicich (2003) allow for structural breaks under the null and the alternative, and the rejection of the null unambiguously implies stationarity.

Lee and Strazicich (2003) consider the model $y_t = \delta' Z_t + X_t$, where $X_t = \beta X_{t-1} + \varepsilon_t$ and Z_t is a vector of exogenous variables. Their LM test has variations for different types of breaks (change in intercept, change in trend slope, and both). For example their model C (change in both intercept and trend) has the following components:

$$Z_t = [1, t, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}]'$$

where $D_{jt} = 1$ for $t \geq T_{B_j+1}$ and 0 otherwise, $DT_{jt} = t - T_{B_j}$ for $t \geq T_{B_j+1}$, and 0 otherwise, $j = 1, 2$. The two break points are denoted as T_{B_1} and T_{B_2}. When the differencing operator is utilized, then $\Delta Z_t = [1, B_{1t}, B_{2t}, D_{1t}, D_{2t}]'$ with $B_{jt} = \Delta D_{jt}$, $D_{jt} = \Delta DT_{jt}$ for $j = 1, 2$.

The test statistic of the two-break LM unit root test is obtained by using the LM (score) principle from the regression:

$$\Delta y_t = \delta' \Delta Z_t + \phi \tilde{S}_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta \tilde{S}_{t-i} + u_t,$$

where $\tilde{S}_{t-1} = y_t - \tilde{\psi}_x - Z_t\tilde{\delta}$, $t = 2, ..., T$ with $\psi_x = y_1 - Z_1\tilde{\delta}$, where $\tilde{\delta}$ are coefficients in the regression of Δy_t on ΔZ_t, and y_1 and Z_1 are the initial observations of y_t and Z_t, respectively. $\Delta \tilde{S}_{t-i}$, $i = 1, ...k$ are augmented terms to correct for autocorrelated errors. The LM-test statistic $\tilde{\tau}$ is the t-statistic testing for $\phi = 0$ (which corresponds to the null hypothesis that a unit root exists).

The optimal number of lags k is determined from Ng and Perron (1995)'s "general to specific" procedure. It starts with $k = 8$. If the last term is not significantly different from zero at the 10% significance level, then the number of lags $k = 7$ is considered and the procedure is repeated again. If the last term significantly differs from zero or $k = 0$, the procedure stops. Practically, at first, the optimal number of lags k is determined for each possible combination of break points. Then, these combinations are examined on the time interval $[0.1T; 0.9T]$. The break dates are determined by the points where the LM t-test statistic is minimized. Critical values are obtained from 20,000 replications of the model with a $T = 100$ sample size.

A.2 Range Unit Root Test and Forward Backward Range Unit Root Test by Aparicio et al. (2006)

Standard unit root tests do not take into account the fact that real macroeconomic data are exposed to structural breaks, outliers and nonlinearity. In such a situation, the power and size of unit root tests can be strongly affected, see e.g. Perron (1990), Perron and Volgelsang (1992) and Perron (2006). The presence of additive outliers affects the size of the test and the null of a unit root can be mistakenly rejected, see Franses and Haldrup (1994). To deal with these problems, Aparicio et al. (2006) suggested the Range unit root (RUR) test, which has a number of advantages: it is invariant to monotonic transformations and model errors, is robust to parameter shifts and structural breaks, and it outperforms standard unit root tests in terms of power when the process is stationary with a near unit root.

For a given time series x_t, Aparicio et al. (2006) consider i-th extremes defined as $x_{1i} = min(x_1, ..., x_i)$ and $x_{ii} = max(x_1, ..., x_i)$ and constructed a sequence of running ranges $R_i^{(x)} = x_{ii} - x_{1i}$, for $i = 1, ...n$, where n is the sample size. To test for the null hypothesis of unit root the following RUR statistic is suggested:

$$RUR \equiv J_0^{(n)} = \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \mathbf{1}(\Delta R_i^{(x)} > 0),$$

where $\mathbf{1}$ () is the indicator function and Δ is the differencing operator.

Under H_0, the test-statistic $J_0^{(n)}$ converges weakly to a non-degenerate unimodal random variable. Under the alternative of stationarity, $J_0^{(n)}$ converges to 0 in probability. Consequently, the left tail of $J_0^{(n)}$ distribution can be used to distinguish between $I(0)$ and $I(1)$ time series without trend, and the right tail for a case of a linear trend.

When additive outliers are considered, Aparicio et al. (2006) proposed an extension of the RUR test known as the Forward-Backward Range unit root (FB-RUR) test. It reduces the size distortion of the test when additive outliers are situated in the beginning of the sample and improves the power compared to the RUR test. For this, the reversed time series $x_t^{'} = x_{n-t+1}$, for $t = 1, ..., n$, are considered, and the analogous sequence of running ranges $R_t^{(x^{'})}$ is constructed as before. The FB RUR test statistic is then as follows:

$$FBRUR \equiv J_*^{(n)} = \frac{1}{\sqrt{2n}} \sum_{i=1}^{n} (\mathbf{1}(\Delta R_i^{(x)} > 0) + \mathbf{1}(\Delta R_i^{(x^{'})} > 0)).$$

Critical values of $J_0^{(n)}$ are obtained from 10,000 replications for different sample sizes and critical levels of the null model with normally distributed errors.

B Periodic Unit Root Tests by Boswijk and Franses (1996) and Franses and Paap (2004)

We report below the case of quarterly data for simplicity, but the results can be extended to monthly data. Consider a Periodic Auto-Regressive (PAR) model of order 1 for a quarterly time series:

$$y_t = \alpha_s y_{t-1} + \varepsilon_t,$$

where ε_t are normally distributed error terms and s corresponds to four seasons, $s = 1, 2, 3, 4$. Denote $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $g(\boldsymbol{\alpha}) = \alpha_1\alpha_2\alpha_3\alpha_4$. The process y_t is called stationary when $|g(\alpha)| < 1$. If $g(\alpha) = 1$, then a unit root exists. We need to test the hypothesis

$$H_0 : g(\alpha) = 1$$

against the alternative $H_1 : |g(\alpha)| < 1$. The unrestricted model is given by

$$y_t = \sum_{s=1}^{4} \alpha_s D_{s,t} y_{t-1} + \varepsilon_t,$$

where $D_{s,t}$ are seasonal dummy variables. When the null condition is imposed, i.e. $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$, the restricted regression is given by

$$y_t = \alpha_1 D_{1,t} y_{t-1} + \alpha_2 D_{2,t} y_{t-1} + \alpha_3 D_{3,t} y_{t-1} + (\alpha_1\alpha_2\alpha_3)^{-1} D_{4,t} y_{t-1} + \varepsilon_t.$$

The parameters α_s can be estimated by ordinary least squares from the unrestricted model and with nonlinear least squares from the restricted model. If we define RSS_0 and RSS_1 as the residual sum of squares for the unrestricted and restricted models, respectively, then the test-statistic of the likelihood ratio test can be computed as

$$LR = n \log \frac{RSS_0}{RSS_1}.$$

Asymptotically, this LR test statistic is distributed as Johansen's trace statistic, and the critical values are tabulated in Osterwald-Lenum (1992), Table 1.1. This test can be generalized to periodic auto-regressions of higher order p, where p is usually determined by using information criteria and checking that the residuals behave approximately as a white noise. For details, see Boswijk and Franses (1996) and Franses and Paap (2004).

If the null hypothesis $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$ of a periodic unit root is not rejected, then we can test two types of parameter restrictions in a second step:

$$\begin{aligned} H_0 &: \alpha_s = 1, s = 1, 2, 3; \\ H_0 &: \alpha_s = -1, s = 1, 2, 3. \end{aligned}$$

If the first H_0 is not rejected, then $\alpha_4 = 1$ and we have a non-periodic unit root, so that the periodic differencing filter can be simplified to $1 - L$, where L is lag operator. If the second H_0 cannot be rejected, then the differencing filter equals $1 + L$ and we have a seasonal unit root. In both cases, the resulting process is a PARI(1), periodically integrated autoregressive model of order 1. Monte-Carlo simulations by Franses and Paap (1994) showed that the maximum-likelihood statistics for testing these H_0 hypotheses follow a standard F-distribution under the null.

C Cointegration Tests allowing for Structural Breaks

C.1 Cointegration test allowing for one break by Gregory and Hansen (1996)

The residual-based cointegration tests by Engle and Granger (1987) and Phillips and Ouliaris (1990) are built without considering any structural break(s). When this is the case, these tests have low power. Gregory and Hansen (1996) proposed a cointegration test which allows for a single endogenous regime shift. The starting point is a model with a standard cointegrating equation (model 1):

$$y_{1t} = \mu + \alpha^T y_{2t} + e_t,$$

where y_{1t} is real-valued, y_{2t} is a $I(1)$ m-dimentional vector and e_t is $I(0)$, $t = 1, ...n$. To develop a model which allows for structural change, a dummy variable is introduced :

$$\phi_{t\tau} = \begin{cases} 0, & if \quad t \leq [n\tau], \\ 1, & if \quad t > [n\tau], \end{cases} ,$$

where $\tau \in (0, 1)$ is a parameter for timing the change point, while the brackets denote the integer part. Gregory and Hansen (1996) considered several specifications, allowing for a change in the intercept μ, in the slope α and with a time trend:

$$y_{1t} = \mu_1 + \mu_2\phi_{t\tau} + \alpha^T y_{2t} + e_t \qquad \text{Level shift (model 2)} \qquad (1)$$

$$y_{1t} = \mu_1 + +\mu_2\phi_{t\tau} + \beta t + \alpha^T y_{2t} + e_t \qquad \text{Level shift with trend (model 3)} \qquad (2)$$

$$y_{1t} = \mu_1 + +\mu_2\phi_{t\tau} + \alpha_1^T y_{2t} + \alpha_2^T \phi_{t\tau} y_{2t} + e_t \qquad \text{Regime shift (model 4).} \qquad (3)$$

(4)

The null hypothesis is no cointegration, while the alternative is cointegration with possible regime shifts (one of the models 2-4). A model is estimated with OLS for each possible τ, and the resulting residuals $\hat{e}_{t\tau}$ are used to compute the first order serial correlation coefficient $\hat{\rho}_\tau$. The second-stage residuals $\hat{v}_{t\tau} = \hat{e}_{t\tau} - \hat{\rho}_\tau \hat{e}_{t-1\tau}$ are then obtained and their long-run variance $\hat{\sigma}_\tau^2$ is estimated (see Gregory and Hansen(1996) for details). The bias-corrected first-order serial correlation coefficient $\hat{\rho}_\tau^*$ is calculated as follows:

$$\hat{\rho}_\tau^* = (\sum_{t=1}^{n-1} \hat{e}_{t\tau}\hat{e}_{t+1\tau} - \hat{\lambda}_\tau)/\sum_{t=1}^{n-1} \hat{e}_{t\tau}^2,$$

where $\hat{\lambda}_\tau$ is the estimation of the weighted sum of autocovariances of $\hat{v}_{t\tau}$. The following test statistics are obtained:

$$\begin{aligned} Z_\alpha(\tau) &= n(\hat{\rho}_\tau^* - 1), \\ Z_t(\tau) &= (\hat{\rho}_\tau^* - 1)/\hat{s}_\tau^2, \end{aligned}$$

where $\hat{s}_\tau^2 = \hat{\sigma}_\tau^2 / \sum_{t=1}^{n-1} \hat{e}_{t\tau}^2$. The third statistic considered by Gregory and Hansen (1996) is the ADF test statistic from an augmented regression of the residuals $\Delta\hat{e}_{t\tau}$ on $\hat{e}_{t-1\tau}$ and lagged first-differences residuals. It is the t-statistic of the regressor $\hat{e}_{t-1\tau}$:

$$ADF(\tau) = tstat(\hat{e}_{t-1\tau}).$$

Since small values of these test statistics are evidence against the null of no cointegration, the smallest values across all possible break points $\tau \in T$ are taken :

$$\begin{aligned} Z_\alpha^* &= \inf_{\tau \in T} Z_\alpha(\tau), \\ Z_t^* &= \inf_{\tau \in T} Z_t(\tau), \\ ADF^* &= \inf_{\tau \in T} ADF(\tau) \end{aligned}$$

Critical values for models 2-4 and different significance levels are provided via simulation methods in Gregory and Hansen (1996).

C.2 Cointegration test allowing for two breaks by Hatemi (2008)

The Hatemi (2008) cointegration test is an extension of Gregory and Hansen's (1996) cointegration test, and it allows for two endogenous breaks. To take two possible changes into account, two dummy variables are introduced:

$$D_{1t} = \begin{cases} 0, & if \quad t \leq [n\tau_1], \\ 1, & if \quad t > [n\tau_1], \end{cases},$$

and

$$D_{2t} = \begin{cases} 0, & if \quad t \leq [n\tau_2], \\ 1, & if \quad t > [n\tau_2], \end{cases},$$

where $\tau_1 \in (0,1)$ and $\tau_2 \in (0,1)$ are unknown parameters for determining the timing of the change points, while the brackets denote the integer part. Considering the case of a level shift, we have the following model:

$$y_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \beta_0' x_t + + u_t.$$

The three residual-based test statistics Z_α, Z_t and ADF are obtained in a similar way to Gregory and Hansen (1996), and their smallest values provide evidence against the null:

$$Z_\alpha^* = \inf_{(\tau_1,\tau_2)\in T} Z_\alpha(\tau_1, \tau_2),$$

$$Z_t^* = \inf_{(\tau_1,\tau_2)\in T} Z_t(\tau_1, \tau_2),$$

$$ADF^* = \inf_{(\tau_1,\tau_2)\in T} ADF(\tau_1, \tau_2).$$

Asymptotic critical values for these tests are obtained by simulation methods and are tabulated in Hatemi (2008).

C.3 Cointegration test by Johansen, Mosconi, and Nielsen (2000) with exogenous structural breaks

The distribution of the standard cointegration test by Johansen (1995) changes if interventional dummies are considered. Johansen et al. (2000) therefore modify this approach, allowing for trend and level breaks at known dates.

Suppose that Y_t is a p-vector process, and that without structural breaks the VECM can be formulated as follows:

$$\Delta Y_t = \Pi Y_{t-1} + \Pi_1 t + \mu + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t,$$

where $\varepsilon_t \sim MN(\mathbf{0}, \Sigma)$. The hypothesis of cointegration can be reformulated as a reduced rank problem of the Π matrix, with $\Pi = \alpha\beta$, where α and β are $(p \times r)$ full rank matrices. In the case that none of the p time series has a quadratic trend, we have that $\Pi_1 = \alpha\gamma'$, where γ is a $(1 \times r)$ full rank matrix.

If there are $q-1$ breaks in the VECM which occur at the dates $T_1, T_2, ..., T_{q-1}$, then the initial sample can be divided into q sample periods $1 = T_0 < T_1 < T_2 < ... < T_q = T$. Here T_j is the last observation of the j-th sub-sample, $j = 1, ..., q$. To account for the break dates, $q-1$ intervention dummy variables are introduced,

$$D_{jt} = \begin{cases} 0, & if \quad T_{j-1}+1 \leq t \leq T_j, \\ 1, & otherwise \end{cases},$$

as well as $q-1$ indicator dummy variables

$$I_{jt} = \begin{cases} 0, & if \quad t = T_{j-1}+1 \\ 1, & otherwise \end{cases}$$

Furthermore, we can define the matrices $D_t = (1, ..., D_{q,t})'$, $\mu = (\mu_1, ..., \mu_q)$ and $\gamma = (\gamma_1', ..., \gamma_q')'$, of dimensions $(q \times 1)$, $(p \times q)$, $(q \times r)$, respectively. The cointegrated VECM model with $q-1$ exogenous breaks can then be written as

$$\Delta Y_t = \alpha \begin{pmatrix} \beta \\ \gamma \end{pmatrix}' \begin{pmatrix} Y_{t-1} \\ tD_{t-k} \end{pmatrix} + \mu D_{t-k} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \sum_{i=0}^{k-1} \sum_{j=2}^{q} \kappa_{j,i} I_{j,t-i} + \varepsilon_t,$$

where $\kappa_{j,i}$ are $(p \times 1)$ vectors and ε_t is a Gaussian white noise vector. For testing the hypothesis of at most r cointegrating relations, the likelihood ratio test statistic is used:

$$LR = -T \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i),$$

where $\hat{\lambda}_i$ are the squared sample canonical correlations of ΔY_t and (Y_{t-1}', tD_{t-k}') corrected for intervention dummies, indicator dummies and augmented terms. Johansen et al. (2000) consider three cases:

1. None of the p time series have a deterministic trend, but the intercepts in the cointegrating relations can vary between sub-samples. The model is denoted $H_c(r)$. The asymptotic distribution of the test-statistic is approximated by a Gamma-distribution.
2. Some or all of the time series follow a trending pattern, the cointegrating relations are stationary in the sub-samples, non-stationary time series and cointegrating relations may have trend breaks. This model is denoted $H_l(r)$.
3. Some or all of the time series follow a trending pattern, the cointegrating relations are stationary in the sub-samples and may have an intercept break, but only non-stationary series may have trend breaks. The model is denoted $H_{lc}(r)$.

The critical values for the models $H_c(r)$, $H_l(r)$ and $H_{lc}(r)$ depend on the number of non-cointegrating relations $p-r$ and the relative location of the breaks. See section 3.4 of Johansen et al. (2000) for more details.

D Periodic Cointegration Tests

Franses and Paap (2004) and Boswijk (1994) suggested a single-equation approach for testing periodic cointegration. For the sake of simplicity, we consider here the bivariate case with two monthly time series $y_{1,t}$ and $y_{2,t}$, where $y_{2,t}$ is weakly exogenous (however, $y_{2,t}$ can easily be extended to a vector of regressors). The conditional periodic error correction model for $y_{1,t}$ is then given by

$$\Delta_{12} y_{1,t} = \gamma_{1s}(y_{1,t-12} - \kappa_s y_{2,t-12}) + \sum_{j=1}^{p-12} \phi_{1s} \Delta_{12} y_{1,t-j} + \sum_{j=0}^{p-12} \phi_{2s} \Delta_{12} y_{2,t-j} + \varepsilon_{1,t},$$

where Δ_{12} is the seasonal difference operator. If needed, the $\Delta_{12} y_{1,t-j}$ and $\Delta_{12} y_{2,t-j}$ variables on the right-hand side can be replaced by $\Delta_1 y_{1,t-j}$ and $\Delta_1 y_{2,t-j}$ variables, see Franses and Paap (2004, p. 111-117) for a discussion. If we define a vector variable w_t which contains the various differenced variables, denote $\delta_{1s} = \gamma_{1s}$, $\delta_{2s} = -\gamma_{1s}\kappa_s$, and $\delta_s = (\delta_{1s}, \delta_{2s})$, then the model can be rewritten as

$$\Delta_{12} y_{1,t} = \sum_{s=1}^{12} (\delta_{1s} D_{s,t} y_{1,t-12} - \delta_{2s} D_{s,t} y_{2,t-12}) + \phi' w_t + \varepsilon_{1,t}.$$

Franses and Paap (2004) extended the Boswijk (1994) cointegration test to test the null hypothesis of no cointegration against the alternative of periodic cointegration and considered two variants: the first one is a Wald test for periodic cointegration in season s, where the null is $\delta_s = 0$ against the alternative of $\delta_s \neq 0$; the second test is a joint Wald test testing the null $\delta_s = 0$ for all seasons s against the alternative that at least one $\delta_s \neq 0$.

Denote $\hat{\delta}_s$ the OLS estimator of δ_s and $\hat{V}(\hat{\delta}_s')$ its estimated covariance matrix. Also, define $\hat{\delta} = (\hat{\delta}_1, \hat{\delta}_2, ..., \hat{\delta}_{12})$ and $\hat{V}(\hat{\delta}')$ its covariance matrix estimator. Then, the two Wald test statistics are given by

$$Wald_s = \hat{\delta}_s' (\hat{V}(\hat{\delta}_s'))^{-1} \hat{\delta}_s,$$

$$Wald = \hat{\delta}' (\hat{V}(\hat{\delta}'))^{-1} \hat{\delta}.$$

Moreover, define RSS_{0s} and RSS_0 as the residual sum of squares under the null for the season-specific and the joint test, respectively, while RSS_1 the residual sum of squares under the alternative for both tests. Let l be the number of estimated parameters. Then, the two test statistics can be re-written as follows:

$$Wald_s = (n-l)\frac{RSS_{0s} - RSS_1}{RSS_1},$$

$$Wald = (n-l)\frac{RSS_0 - RSS_1}{RSS_1}.$$

The previous tests can be easily extended to the case involving seasonal intercepts and trends, see Franses and Paap (2004) for details. The asymptotic distribution of the test statistics is non-standard and critical values are reported in Table C.1 in Franses and Paap (2004).

The previous test for periodic cointegration assumes that $y_{2,t}$ is weakly exogenous. To test this hypothesis, Boswijk (1994) suggests adding the periodic error terms $D_{s,t}(y_{1,t-12} - \hat{k}_s y_{2,t-12})$ to an autoregressive model for $\Delta_{12} y_{2,t}$ which also includes lags of $\Delta_{12} y_{1,t}$. Under the null of weak exogeneity, the Likelihood Ratio test that the parameters of the periodic error terms are zero for all s is asymptotically $\chi^2(12)$ distributed. When the null hypothesis of weak exogeneity is rejected, alternative methods have to be used (for example, dynamic-OLS; see Boswijk and Franses (1995) for details).

E Bayesian VARs

Bayesian methods treat the true value of the unknown parameter vector θ as a probability distribution $\pi(\theta|y)$, which is the called posterior distribution of θ given data y. The prior distribution, $\pi(\theta)$, is set externally and reflects the researcher's beliefs on the unknown parameter of interest, while $l(y|\theta)$ is the likelihood distribution, which depends on the information from the given data y. Bayes' theorem links all these distributions through this formula:

$$\pi(\theta|y) = \frac{\pi(\theta) l(y|\theta)}{\int \pi(\theta) l(y|\theta) d\theta}.$$

Given that the denominator is a normalizing constant, the posterior is proportional to the product of the likelihood and the prior, i.e. $\pi(\theta|y) \propto \pi(\theta) l(y|\theta)$.

Consider the following VAR model of order p for the m-dimensional vector y_t:

$$y_t = a_0 + \sum_{j=1}^{p} A_j y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d.N(0, \Sigma_\varepsilon)$$

for $t = 1, \ldots, T$, where ε_t is an error vector term. In matrix-vector notation, it takes the form

$$y = (I_m \otimes X)\theta + e,$$

where I_m is an $m \times m$ identity matrix, $X = (x_1, \ldots, x_t)'$ is a $T \times (mp+1)$ matrix for $x_t = (1, y'_{t-1}, \ldots, y'_{t-q})$, $\theta = vec(A)$, and $e \sim N(0, \Sigma_\varepsilon \otimes I_T)$.

In this work, we used the Litterman/Minnesota prior, which is a family of priors where Σ_ε is known and replaced with an estimated $\hat{\Sigma}_\varepsilon$. The Minnesota prior assumes that $\theta \sim N(\theta_0, V_0)$, where $\theta_0 = \mu_1 \cdot j_{mp}$, j_{mp} is an mp-element unit vector, and $\mu_1 = 0$ is a hyper-parameter. V_0 is a non-zero covariance matrix constructed as follows: the elements of V_0 which correspond to exogenous variables are set to infinity, and the remaining part is a diagonal matrix with the following diagonal elements:

$$v_{ij}^l = \begin{cases} \left(\frac{\lambda_1}{l^{\lambda_3}}\right)^2 & \text{for} \quad i = j \\ \left(\frac{\lambda_1 \lambda_2 \sigma_i}{l^{\lambda_3} \sigma_j}\right)^2 & \text{for} \quad i \neq j \end{cases},$$

where $l = 1, \ldots, p$ and σ_i is the i-th diagonal element of $\hat{\Sigma}_\varepsilon$. The scalars λ_1, λ_2, λ_3 control the overall tightness, relative cross-variable weight and the decay of lag coefficients, respectively. We chose $\lambda_1 = 0.1$, $\lambda_2 = 1$, $\lambda_3 = 1$. Given the Minnesota prior, the posterior distribution of the parameter θ is given by

$$\theta \sim N(\bar{\theta}, \bar{V}),$$

where

$$\bar{V} = \left(V_0^{-1} + \left(\hat{\Sigma}_\varepsilon^{-1} \otimes X'X\right)\right)^{-1},$$

$$\bar{\theta} = \bar{V}\left(V_0^{-1}\theta_0 + \left(\hat{\Sigma}_\varepsilon^{-1} \otimes X\right)' y\right).$$

F The Model Confidence Set

The Model Confidence Set (MCS) approach by Hansen et al. (2011) can be used to select the best forecasting models among a set of models, given a confidence level α.

First, the MCS procedure applies an equivalence test δ_M to the set of forecasting models $M = M_0$, to test the null hypothesis of equal forecasting accuracy,

$$H_{0,M} = E(d_{ij,t}) = 0, \quad \forall i, j \in M,$$

where $d_{ij,t} = L_{i,t} - L_{j,t}$ is the sample loss differential between forecasting models i and j and $L_{i,t}$ stands for the loss function of model i at time t. The alternative hypothesis $H_{A,M}$ is that $E(d_{ij,t}) \neq 0$ for some $i, j \in M$. If the null cannot be rejected, then $\widehat{M}^*_{1-\alpha} = M$. When the null is rejected, it indicates that some of the models of the set M have worse sample performance than others. Therefore, the elimination rule e_M is used to remove these models from the set M. The procedure is repeated until the null cannot be rejected, and the resulting models define the model confidence set $\widehat{M}^*_{1-\alpha}$.

Hansen et al. (2011) proposes different equivalence tests and we discuss here the Semi-Quadratic statistic which we used in the paper. First, the following t-statistics are computed:

$$t_{ij} = \frac{\overline{d}_{ij}}{\widehat{var}(\overline{d}_{ij})}, \quad \text{for} \quad i, j \in M,$$

with $\overline{d}_{ij} = T^{-1}\sum_{t=1}^{T} d_{ij,t}$. Then, the semi-quadratic statistic, $T_{S,Q}$, is computed as follows:

$$T_{S,Q} = \sum_{i,j \in M} t_{ij}^2.$$

The distribution of this test statistic is non-standard and is estimated using bootstrapping methods, see Hansen et al. (2011) for details. For all tests, the same significance level α is used, which asymptotically guarantees that $\Pr(M^* \subset \widehat{M}^*_{1-\alpha}) \geq 1-\alpha$, where $\widehat{M^*}$ is the set of models with a given confidence level. If only one model is included in M^*, we have $\lim_{n\to\infty} \Pr(M^* = \widehat{M}^*_{1-\alpha}) = 1$.

G Nonlinear Models

We considered three nonlinear models: the first one was the SETAR model with 2 regimes:

$$Y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}Y_{t-1} + \ldots + \phi_{1,p}Y_{t-p} + \varepsilon_t, & \text{if} \quad Y_{t-1} \leq c \\ \phi_{0,2} + \phi_{1,2}Y_{t-1} + \ldots + \phi_{1,p}Y_{t-p} + \varepsilon_t, & \text{if} \quad Y_{t-1} > c \end{cases}$$

where c is a threshold to be estimated and which identifies the two regimes. We allowed the number of lags p to vary between 1 and 12, while Y_t was either the log-prices or the log-returns, for a total of 24 models.

The second nonlinear model was the logistic smooth transition autoregressive (LSTAR) model, which is a generalization of the SETAR model:

$$\begin{aligned} Y_t \quad = \quad & (\phi_{0,1} + \phi_{1,1}Y_{t-1} + \ldots + \phi_{1,p}Y_{t-p})[1 - G(Y_{t-1}, \gamma, c)] + \\ & +(\phi_{0,2} + \phi_{1,2}Y_{t-1} + \ldots + \phi_{1,p}Y_{t-p})[G(Y_{t-1}, \gamma, c)] + \varepsilon_t \end{aligned}$$

where $G(Y_{t-1}, \gamma, c) = [1 + \exp(-\gamma(Y_{t-1} - c))]^{-1}$ is the first order logistic transition function, bounded between 0 and 1, γ is the slope parameter and c is the location parameter. Differently from SETAR models, the LSTAR model assume that the change between the two regimes is gradual and smooth, see Tong (1990) for a discussion at the textbook level. We allowed again the number of lags p to vary between 1 and 12, while Y_t was either the log-prices or the log-returns, for a total of additional 24 nonlinear models.

Finally, we considered the additive autoregressive model (AAR), also known as generalized additive model (GAM), since it combines generalized linear models and additive models:

$$Y_t = \phi_0 + s_1(Y_{t-1}) + \ldots + s_p(Y_{t-p}) + \varepsilon_t$$

where s_i are smooth functions represented by penalized cubic regression splines, see Wood (2006) for a discussion at the textbook level. The number of lags p varied between 1 and 12, while Y_t was either the log-prices or the log-returns, for a total of additional 24 models.