8th International Congress on Advances in Civil Engineering, 15-17 September 2008, Eastern Mediterranean University, Famagusta, North Cyprus

Spreadsheet Modelling Approach for Estimating Car Ownership

O. Baskan, H. Ceylan, S. Haldenbilen, H. Ceylan

Department of Civil Engineering, Engineering Faculty, Pamukkale University, Denizli, Turkey

Abstract

This study develops a Spreadsheet Solver Model (SSM) to determine car ownership figures based on vehicular credits, fuel prices and gross domestic product per capita (GDPPC). The monthly variations on vehicular selling are also included in the SSM. The definition of the SSM and application is correspondingly defined on the spreadsheet. The effect of inflation rate is included to estimate the amount of vehicular credits and hence motor vehicle sales. Vehicle ownership figures are obtained under four scenarios. Results showed that the SSM can be used for obtaining car ownership in future and that monthly changes may be estimated, and then car manufacturers and tax makers will manage the vehicular business according to the SSM results.

Keywords: *Car ownership, vehicle ownership, vehicle credit, GDPPC, spreadsheet.*

1. Introduction

Vehicular business cycle studies play an important role in the decision-making processes of government agencies and private sectors. An increase on vehicular business on any country will cause to increase on mobility need for people. Mobility requirement lead to a human decision-making process to which of the options (transport mode) will solve his or her problem in a quickest way. The people then may usually prefer to buy a car, especially for short distance travelling. Owing a car may be also considered as a status symbol for society. Decisions to buy a car with scarce economic sources usually require using finance provided by banks. The income is also another important parameter for

decision-making. It may then be stated that the increase on the amount of vehicular credit may lead to increase on car ownership and hence rapid increase on traffic problems.

Lam and Tam (2002) presented an aggregate car ownership model for Hong-Kong and examined the reliability of territory-wide car ownership figures. The effect of income distribution was studied by Dargay (2001). It was found that the elasticity with respect to rising income was significantly greater than the elasticity with respect to the falling income. Dargay and Gately (1999) projected the car growth and total vehicle stock until 2015 for OECD countries, China, India, and Pakistan. Ben-Akiva et al. (1981); Janson (1989) Pendyala et al. (1995) and Kumar and Krishna (2006) used socio-economic variables to determine car ownership figures, but they may well be extended to include the economic stability which can be measured by a vehicular credit system in a country. Given the importance of car ownership in transport, land-use planning and its relationship with sectoral energy consumption, the environment and health, it is the objective of this research to develop car ownership models and to apply the models into the policy making scenarios. The paper is organized as follows. The model development is presented in the next section. Data collection on socio-economic indicators is described in Section 3. The application of the SSM is carried out in Section 4. Section 5 is on the future projection. Finally, the Section 6 presents the conclusions.

2. Model Development

Spreadsheet tools are becoming recently increasingly popular for researchers. The popularity is because of fast and easy development of both linear and non-linear models using spreadsheet tools. All the major spreadsheet packages include an optimization tool known as Solver. It can be used to maximize or minimize the value of a "target" worksheet cell by altering the values of other selected "changing" cells in the spreadsheet that influence the value in the target cell. Solver also allows constraints to be placed on the values of any cells in the worksheet.

There are many situations in which a dependent variable moves in a nonlinear fashion with dependent variables. In these cases, it is usually tried to identify a nonlinear function $F(x)$ such that $F(x) = f(x) + \varepsilon_x$, where ε_x represents a random disturbance term. Because $f(x)$ is nonlinear, it is usually not possible to use linear regression to estimate the parameters involved in $f(x)$. As in car ownerships model, the SSM is formed as:

$$F(x) = (w_0 * {X_1}^{w_1} + w_2 * {X_2}^{w_3} + {X_3}^{w_4}) * S_i + \varepsilon_x \qquad i = 1, 2, \ldots, 12 \qquad (1)$$

Where, $F(x)$ is the number of sold motor vehicles, X_1 is the amount of vehicle credits (10^6\$), X_2 is the fuel price (\$/lt), X_3 is the GDPPC (\$), S_i is the monthly index and w_0, w_1, w_2, w_3, w_4 are the coefficients and ε_x is given as zero. The solution of Eqn. (1) is given in the next section. During the model development

stage, we first need to find a correlation among the selected variables. The highest correlation is between number of sold vehicles and vehicle credits (0.85), and the lowest one is the fuel price (0.45). The increase on fuel price may not considerably affect the number of sold vehicles since there is not a significant correlation among independent variables. During the SSM validation, two-third of the data are used for model building and the one-third of the data is used for testing, where minimum sum of squared error (SSE) are used to solve Eqn. (1).

3. Data Collection

The numbers of monthly sold motor vehicles for the period of 1996-2004 are collected from National Statistics (NS, 2005). There were also lowest vehicle sales in 2002 since there were economic crises at the end of year 2001 in Turkey. After that the economical stability were established and the number of sold vehicles have increased and it reached to peak level of 120 000 in December in 2004. The amount of the vehicular credits is obtained from The Banks Association of Turkey (BAT) (2005). The amount of vehicular credits was reached to $990*10^6$ ($) at the end of 2004.

Fuel prices which are collected from the NS (2005). There is a continuous increase on the fuel prices for the period of 1996-2004. Whilst fuel prices were about $ 0.60/lt in 1996, then it steadily increased to $ 1.60/lt in 2004. Data for the GDPPC are collected from Central Bank of Turkey (CBT, 2005). Monthly GDPPC is obtained by dividing the annual values to 12 months. The effects of economic crises which were occurred at the end of year 2001 can be seen an exact decrease for the GDPPC. The observed monthly inflation rates are obtained from the CBT (2005) for the period of 1996-2004. There is a tendency of decreasing the inflation although there is a big fluctuation in 1999.

4. Model Application

In order to find the number of sold vehicles, the regression model (RM) is carried out first as:

$$f(x) = w_0 + w_1 * X_1 + w_2 * X_2 + w_3 * X_3 \qquad (2)$$

Where $f(x)$ is the number of sold motor vehicles, X_1 is the amount of vehicle credit (10^6 $), X_2 is the fuel price ($/lt), X_3 is the GDPPC ($) and w_0, w_1, w_2, w_3 are the coefficients.

The solution of the model is:

$$f(x) = 23790.09 + 81.42 * X_1 - 38897.50 * X_2 + 121.71 * X_3 \quad R^2 = 0.66 \qquad (3)$$

During the RM analysis, 108 of collected data are used for model development and 12 are used for testing. The performance of Eqn. (3) is compared with the observed values and SSM. In testing period, total absolute average relative error

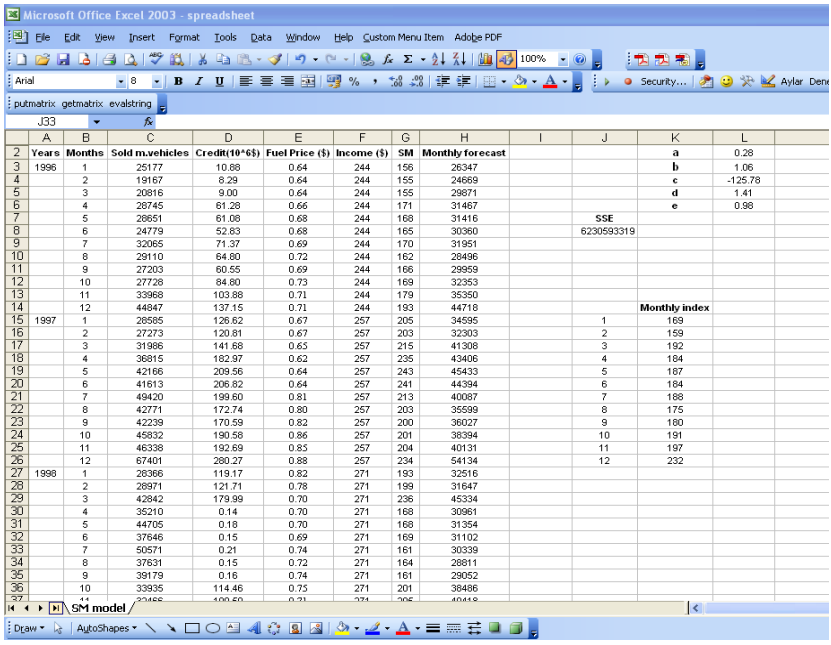
is about 14% for RM. The solution of the SSM model is carried out on the spreadsheet with solver where all the data are input to the cells. Part of the data and solution of the SSM can be seen in Figure 1. The key cell formulas and solution procedure is given in Table 1. The solution of the SSM with spreadsheet is given in Eqn. (4) as:

$$F(x) = (0.28 * X_1^{1.06} - 125.78 * X_2^{1.41} + X_3^{0.98}) * S_i \quad (4)$$

where the optimum values of S_i can be seen on the K[th] column from 15[th] to 26[th] row in Figure 1. The performance of the SSM is compared with the observed values and the RM and results are given in Table 2. In testing period, total absolute average relative error is 8.79%. The performance of the SSM is better than the RM for this study.

Table 1. Key cell formulas

Cell	Formula	Copied to
G3	==$L$2*D3^$L$3+$L$4*E3^$L$5+F3^$L$6	G3:G302
K15	=SUMIF($B$3:$B$110;J15;$G$3:$G$110)/COUNTIF($B$3:$B$110;J15)	K15:K26
H3	=G3*VLOOKUP(B3;$J$15:$K$26;2)	H3:H302
J8	=SUMXMY2(H3:H110;C3:C110)	



	A	B	C	D	E	F	G	H	I	J	K	L
2	Years	Months	Sold m.vehicles	Credit(10^6$)	Fuel Price ($)	Income ($)	SM	Monthly forecast			a	0.28
3	1996	1	25177	10.88	0.64	244	156	26347			b	1.06
4		2	19167	8.29	0.64	244	155	24669			c	-125.78
5		3	20816	9.00	0.64	244	155	29871			d	1.41
6		4	28745	61.28	0.66	244	171	31467			e	0.98
7		5	28651	61.08	0.68	244	168	31416		SSE		
8		6	24779	52.83	0.68	244	165	30360		6230593319		
9		7	32065	71.37	0.69	244	170	31951				
10		8	29110	64.80	0.72	244	162	28496				
11		9	27203	60.55	0.69	244	166	29959				
12		10	27728	84.80	0.73	244	169	32353				
13		11	33968	103.88	0.71	244	179	35350				
14		12	44847	137.15	0.71	244	193	44718			Monthly index	
15	1997	1	28585	126.62	0.67	257	205	34595		1	169	
16		2	27273	120.81	0.67	257	203	32303		2	159	
17		3	31986	141.68	0.65	257	215	41308		3	192	
18		4	36815	182.97	0.62	257	235	43406		4	184	
19		5	42166	269.56	0.64	257	243	45433		5	187	
20		6	41613	206.82	0.64	257	241	44394		6	184	
21		7	49420	199.60	0.81	257	213	40087		7	188	
22		8	42771	172.74	0.80	257	203	35599		8	175	
23		9	42239	170.59	0.82	257	200	36027		9	180	
24		10	45832	190.58	0.86	257	201	38394		10	191	
25		11	46338	192.69	0.85	257	204	40131		11	197	
26		12	67401	280.27	0.88	257	234	54134		12	232	
27	1998	1	28366	119.17	0.82	271	193	32516				
28		2	28971	121.71	0.78	271	199	31647				
29		3	42842	179.99	0.70	271	236	45334				
30		4	35210	0.14	0.70	271	168	30961				
31		5	44705	0.18	0.70	271	168	31354				
32		6	37646	0.15	0.69	271	169	31102				
33		7	50571	0.21	0.74	271	161	30339				
34		8	37631	0.15	0.72	271	164	28811				
35		9	39179	0.16	0.74	271	161	29052				
36		10	33935	114.46	0.75	271	201	38486				

Figure 1. Spreadsheet for the SSM and indexing

Table 2. Comparison of performance of the RM and SSM

Years	Observed	RM	SSM	Relative error of RM (%)	Relative error of SSM (%)
Jan-04	49 560	57 132	56 147	-15.28	-13.29
Feb-04	33 398	42 891	41 211	-28.42	-23.39
Mar-04	77 948	79 097	86 435	-1.47	-10.89
Apr-04	82 221	68 686	72 546	16.46	11.77
May-04	74 638	67 428	72 046	9.66	3.47
Jun-04	61 953	57 882	61 881	6.57	0.12
Jul-04	82 597	74 989	79 092	9.21	4.24
Aug-04	67 111	61 241	61 544	8.75	8.30
Sep-04	64 291	60 282	62 421	6.24	2.91
Oct-04	69 254	58 137	64 063	16.05	7.50
Nov-04	88 450	67 539	75 253	23.64	14.92
Dec-04	115 963	85 914	110 505	25.91	4.71
Average relative error (%)				13.97	8.79

5. Forecasting Motor Vehicle Sales

The projection is made until 2020 since the longer period may not be reasonable due to the country's European direction.

5.1. Vehicle credits

In order to forecast the vehicle credits in future, it is required to estimate inflation rate since interest rates of vehicular credits differ because of inflation. It is estimated with the following non-linear time series model which is fitted to the observed inflation for the period of 1996-2004.

$$y_{\inf} = 8.97 * t^{-0.34} \qquad R^2=0.60 \tag{5}$$

Where $y_{\inf}$ is the inflation rate and t is a time series for which 1996=1, 1997=2,…,2004=9

After determining inflation rate, estimation of vehicle credits is obtained with a time-lagged time series model as:

$$VC_n = w_0 * VC_{n-2} + w_1 * VC_{n-1} + w_2 * y_{\inf,n} \tag{6}$$

Where VC is vehicular credit amount (10^6\$), w_0, w_1, w_2 are the coefficients and n is the index of time series with reference month n for which 1996=1, 1997=2,……., 2004=9. The solution of Eqn. (6) with solver is given in Eqn. (7).

$$VC_n = 0.335 * VC_{n-2} + 0.675 * VC_{n-1} - 0.08 * y_{\inf,n} \qquad R^2=0.83 \tag{7}$$

5.2. Fuel prices

Expected fuel prices are determined until 2020 with exponential and polynomial time series in the following way. The exponential form is given in Eqn. (8).

$$FP = 0.3956 * e^{0.0076x} \qquad R^2 = 0.79 \qquad (8)$$

The polynomial form is:

$$FP = 4.1*10^{-5} * x^2 - 0.0029 * x + 0.65 \qquad R^2 = 0.83 \qquad (9)$$

Where *FP* is the fuel price ($/lt), *x* is the time series for which 1996=1, 1997=2,……,2004=9. The expected prices of fuel will be between 4.00 $/lt and 4.50 $/lt in December in 2020.

5.3. GDPPC

The GDPPC is estimated as:

$$Y = 1268.5 * e^{0.0423x} \qquad R^2 = 0.74 \qquad (10)$$

Where *y* is the GDPPC ($10^6$$), *x* is the time series for which 1996=1, 1997=2,…….2004=9. It will reach to a level of 600 $/month in 2020.

Car ownership projection is made under four scenarios since it will reflect the country's economic development. The average results of four scenarios are given in Figure 2. The proposed scenarios are the followings.

Scenario 1: Fuel price is high and GDPPC is low;
Scenario 2: Fuel price is low and GDPPC is low;
Scenario 3: Fuel price is high and GDPPC is high; and
Scenario 4: Fuel price is low and GDPPC is high.

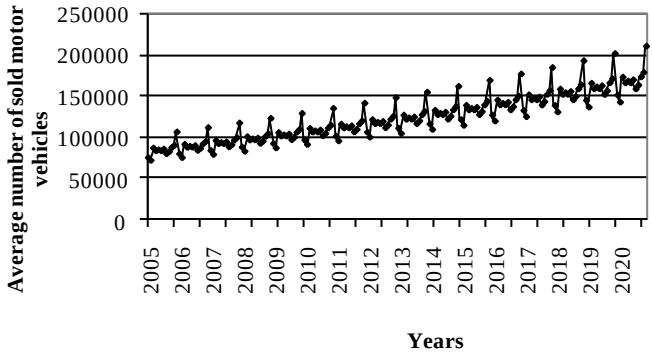


Figure 2. Forecasted average number of sold motor vehicles under four cases

The highest motor vehicle selling is at the end of each year. This is due to the tax damping on first hand vehicles and the car manufacturers who wants to sell all the stocks before any new year starts. After obtaining monthly motor sale, the relationship between motor vehicle ownership and yearly sold motor vehicles are obtained as:

$$y = 338760 * \ln(x) - 1E + 06 \qquad R^2 = 0.82 \qquad (11)$$

where y is the motor vehicle ownership and x is the total yearly sold motor vehicles. Figure 3 shows relationship between the motor vehicle ownership and motor vehicle sale.

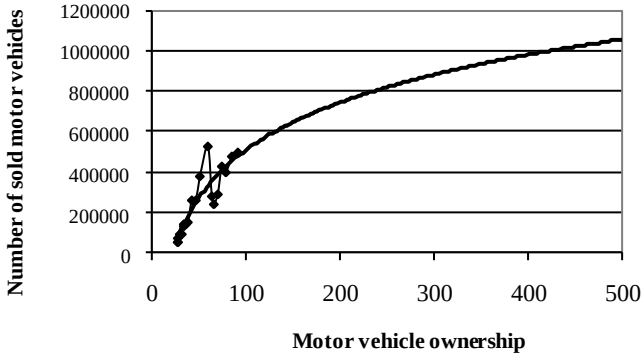


Figure 3. Relationship between motor vehicle ownership and yearly sold motor vehicles

Motor vehicle ownership is estimated until 2020 to determine the ratio of the car ownership out of total number of motor vehicles. Car ratios are obtained using Equation (12) as:

$$y = 0.01 * \ln(x) + 0.37 \qquad R^2 = 0.87 \qquad (12)$$

where y is the ratio of car of total number of motor vehicles and x=1 for 1965, x=2 for 1966,….x=39 for 2004. Estimated car ownership figures that are obtained using from Equation (11) and (12), are given in Table 3 under four scenarios. An average car ownership figures will vary from 256 to 296 per thousand populations.

Table 3. Car ownership for per thousand

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Years	Car ownership	Car ownership	Car ownership	Car ownership
2005	88	88	88	89
2010	136	138	142	145
2015	192	197	207	212
2020	256	270	282	296

6. Conclusions

This study obtains car ownership figures based on the number of motor vehicles, vehicular credits, GDPPC and fuel price. The monthly changes of the vehicle sale are obtained with the SSM. The definition of the model is given and the SSM application is correspondingly outlined. Projection is made until 2020 under four scenarios. The following conclusions can be drawn from this study.

There is a highest correlation between vehicle credit and number of sold vehicles. The vehicular business will also affected by the fuel price. There is no effect is obtained with inflation rate and vehicles sold, but it affects the vehicular credit.
The RM and the SSM are developed and the total absolute error difference between two models is about %5. Thus, the SSM may be used for future projections. An average car ownership will reach to a level of 276 per thousand in 2020. It is obtained that monthly changes may be estimated, and then car manufacturers and tax makers will manage the vehicular business according to the SSM results. The planners may also get benefited from the number of vehicles sold in terms of how the car ownership will grow in future. The vehicular tax income rate from vehicular business may easily be estimated by the national budget makers. This study shows that the determination of vehicular business with monthly bases may easily be carried out with spreadsheets. The relation between the monthly vehicle sale and accident occurrence might be analyzed and this would provide an useful results for engineers and policy makers. The future work should be on transforming vehicular sale to a monthly change on car ownership and then the accident relation studies.

Acknowledgement

This research is sponsored by the Scientific & Technological Research Council of Turkey (TUBITAK) with the project number 104I119.

References

BAT, The Banks Association of Turkey. (2005). Downloadable from http://www.tbb.org.tr/english

Ben-Akiva, M., Manski, C. F. , Sherman, L. A. (1981). Behavioral approach to modelling household motor vehicle ownership and applications to aggregate policy analysis. *Environment and Planning*, 13(A), 399-411.

CBT, Central Bank of Turkey. (2005). Downloadable from http://www.tcmb.gov.tr

Dargay, M. J. (2001). The effect of income on car ownership: evidence of asymmetry. *Transport Research Part A*, 35, 807-821.

Dargay, J., Gately, D. (1997). Vehicle ownership to 2015: implications for energy use and emissions. *Energy Policy*, 25, 1121-1127.

Dargay, J., Gately, D. (1999). Income's effect on car and vehicle ownership, worldwide: 1960-2015. *Transport Research Part A*, 33, 101-138.

Haldenbilen, S., Ceylan, H. (2005a). Transport demand estimation based on genetic algorithm approach. *Transportation Planning and Technology*, Vol.28(6), 403-426.

Haldenbilen, S., Ceylan, H. (2005b). Genetic algorithm approach to estimate transport energy demand in Turkey. *Energy Policy*, 33, 89-98.

Janson, J. O. (1989). Car demand modeling and forecasting: a new approach. *Journal of Transport Economics and Policy*, 23, 125-140.

Kumar, M., Krishna Rao, K.V. (2006). A Stated Preference Study for a Car Ownership Model in the Context of Developing Countries. *Transportation Planning and Technology*, Vol. 29(5), 409-425.

Lam, W., Tam, M. (2002). Reliability of territory-wide car ownership estimates in Hong-Kong. *Journal of Transport Geography*, 10, 51-60.

NS, National Statistics. (2005). Downloadable from http://www.die.gov.tr.

Ogut, K.S. (2006). Modelling Car Ownership in Turkey Using Fuzzy Regression. *Transportation Planning and Technology*, Vol. 29(3), 233-248.

Pendyala, R. M., Kostyniuk, L. P., Goulias, K. G. (1995). A repeated cross-sectional evaluation of car ownership. *Transportation* , 22, 165-184.

Ragsdale, C.T., Plane, D.R. (1999). On modelling time series data using spreadsheets. *Omega*, 28, 215-221.

Shantanu, B., Gary E. B. (2000). Use of a network model interface to build spreadsheet models of process systems: a productivity enhancement tool for risk management studies. *Computers and Chemical Engineering*, 24, 1511-1515.