Journal of Business Research 67 (2014) 1226–1232

Contents lists available at ScienceDirect

Journal of Business Research

Using diffusion models to forecast market size in emerging markets with applications to the Chinese car market☆

Lixian Qian [a], Didier Soopramanien [b,*]

[a] Business School, Xi'an Jiaotong–Liverpool University, 111 Ren'ai Road, Suzhou Industrial Park, Suzhou, Jiangsu Province 215123, China
[b] Department of Management Science, Lancaster University Management School, Lancaster LA1 4YX, UK

ARTICLE INFO

Article history:
Received 1 September 2011
Received in revised form 1 April 2013
Accepted 1 April 2013
Available online 22 May 2013

Keywords:
Market size forecasting
Diffusion model
Chinese car market
Market composition change

ABSTRACT

Marketing managers have to forecast the market size and this forecast guides strategic decisions whether to continue exporting, open new factories or expand existing production operations. Forecasting sales and the market size is a challenging task; even more so in emerging markets where data is limited and the market demand is changeable. This research proposes a novel approach that applies diffusion models using car ownership data to forecast car sales. Car ownership data may be easier to access than sales data in emerging markets but marketing managers are more interested in the sales forecast. Researchers propose using diffusion models to forecast the adoption of new products or products which are new to consumers in a market. This research demonstrates that marketing managers can use diffusion models to predict car sales in China where cars are new products to most consumers in this market. Since the majority of car buyers in China are first time buyers, car manufacturers and retailers must also forecast when the market composition will change. This effectively means predicting when first time car buying will start to slow down and repeat/replacement purchase or second hand car purchase will become more important. To forecast both sales and market composition change, marketing managers must choose a robust model. Managers want insights from models that have been tested robustly especially in less stable market conditions. In this context, this study illustrates the value of using a rolling forecast instead of a fixed horizon approach when comparing and choosing which model to use to forecast both sales and market composition change for the Chinese car market.

1. Introduction

Forecasting the market size for any product or service informs strategic and investment decisions. The market size forecast influences strategic decisions of firms whether they should enter new markets or if they should expand existing capacity in markets where they already operate. Multinational car companies that are interested in manufacturing and selling cars in China also have to make such strategic decisions. Decision makers, in such instances, want to use data and insights that they can rely upon with confidence. This is particularly important when the market conditions are changing rapidly.

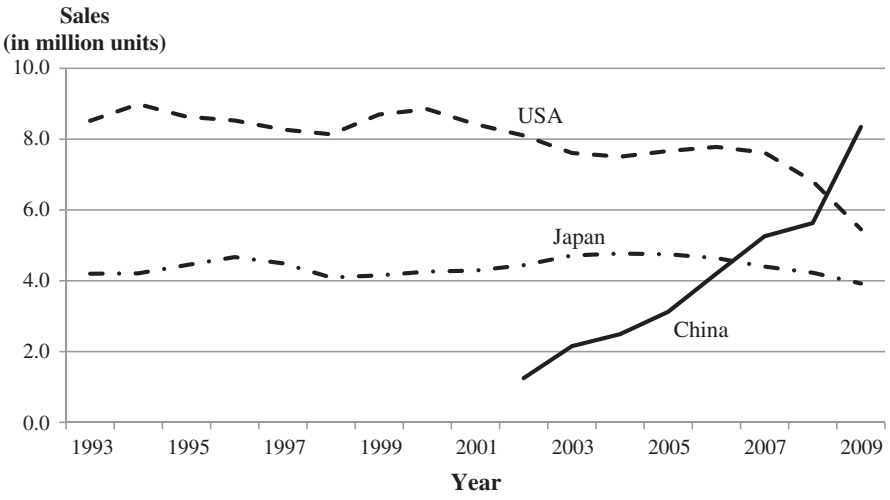
The Chinese car market provides an interesting context for this research, as it is an example of a typical market in an emerging economy. The Chinese car market has experienced exponential growth in the last decade compared to other more mature markets. In 2010, China overtook the United States to become the largest car market in the world when comparing annual car sales. Annual sales of new cars in China, the United States and Japan are shown in Fig. 1. This clearly depicts a non-linear growth pattern of car sales in China. The Chinese car market is far from reaching market saturation where growth slows down and sales/car ownership reaches a plateau. The average car ownership has increased from 90 cars per 1000 inhabitants in the early1990s to 118 cars per 1000 inhabitants in 2005 (World Development Indicators published by World Bank (see Table 1)). Car markets in more mature economies, represented by the US, UK, Germany and Japan, have high car ownership rate with approximately 500 cars per 1000 inhabitants in 2005. China has a much lower car ownership level: every 1000 inhabitants only owned 15 cars in 2005 but this statistic has increased by more than 10 times since 1990.

Researchers propose using diffusion models when data is limited especially in emerging markets because such types of models require limited information (Dargay & Gately, 1999; De Jong, Fox, Daly, Pieters, & Smit, 2004). Previous research advocating the use of diffusion models focus on forecasting ownership level of new products. This study proposes an approach to forecast car sales using ownership data that can be employed when first time buyers dominate the market, which is the case for the Chinese car market. Meade and Islam (2006) recommend comparing different specifications of diffusion models and how they are used to forecast sales. Fildes, Nikolopoulos, Crone, and Syntetos (2008)

☆ The authors thank Robert Fildes and Christopher Kirkbride (both from Lancaster University Management School) for their comments on earlier versions of the paper. The authors would also like to thank the reviewers and the editor for their comments on the paper.

* Corresponding author. Tel.: +44 1524 593874; fax: +44 1524 844885.
E-mail addresses: lixian.qian@xjtlu.edu.cn (L. Qian), d.soopramanien@lancaster.ac.uk (D. Soopramanien).

0148-2963/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jbusres.2013.04.008



Data Sources: China Associate of Automobile Manufacturers (CAAM), US National Automobile Dealers Association (NADA) and Japan Automobile Manufacturers Association (JAMA)

Fig. 1. New car sales history in China, USA and Japan.

argue that researchers must justify the choice of a forecasting model by conducting robust comparison of the proposed models. This research considers five different specifications of diffusion models and compares the sales forecasts generated from these models to those generated by three benchmark models that employ sales data. The benchmark models are exponential smoothing, ARIMA, and linear regression model.

This study also demonstrates the importance of using a rolling forecast instead of a fixed horizon for the validation sample to compare the models especially when the market conditions are changing rapidly. If a researcher chooses a particular model based on the forecasting performance over a fixed horizon, it is possible that he/she may come to a different conclusion if the researcher changes the length of that horizon (Fildes, 1992; Tashman, 2000). The rolling forecast methodology is, in other words, more robust and employing this approach is particularly valuable when market conditions are changeable as is the case for the Chinese car market.

Researchers typically use diffusion models to forecast the adoption of a new product by first time buyers (Bass, 1969; Meade & Islam, 2006; Ratchford, Balasubramanian, & Kamakura, 2000). This study investigates whether such types of models can be used to predict when first time car purchase rate is going to slow down in China. This insight has important practical implications for car manufacturers and retailers because when first time car buying starts to slow down, they will have to change their marketing strategy to target those first time buyers who will replace their existing cars and consumers who may want to buy second hand cars.

The remainder of this paper has the following organisation. Section 2 briefly describes the specifications and applications of diffusion models. Section 3 presents the empirical approach and the data followed by the estimated models in section 4. In section 5 we evaluate the sales forecasting performance of the models. In the last section we summarize and comment on the findings of the research.

Table 1
Number of cars per 1000 people worldwide and across some countries.
Data source: 2006 World Development Indicators, World Bank.

Year	World	China	India	Brazil	Germany	Japan	UK	US
1990	91.06	1.43	2.42	NA	386.30	283.32	340.58	573.28
1995	89.65	2.90	4.12	119.73[b]	494.90	356.19	352.02	484.84
2000	104.35	7.00	6.00	134.00	515.00	413.00	388.00	473.00
2005	117.91[a]	15.00	8.00[c]	136.00[a]	550.00	441.00[a]	457.00	461.00[d]

[a] Data in 2004.
[b] Data in 1996.
[c] Data in 2003.
[d] Excludes personal passenger vans, passenger minivans, and utility-type vehicles, which are all treated as trucks.

2. Diffusion models and their applications

De Jong et al. (2004) propose using diffusion models to forecast the car ownership level where the explanatory variables include a trend variable and/or other macroeconomic variables such as GDP per capita. Diffusion models and the resulting diffusion curve typically assume that the diffusion rate of a product within a population follows a sigmoid shape. That is, at the early stage of adoption, growth rate is slow and after crossing a threshold, adoption of the new product rises rapidly until it approaches the saturation level. In this section we briefly describe three well-known diffusion models: Gompertz, Logistics and Bass diffusion models. We also highlight how these models have been applied to forecast car ownership level. For a more comprehensive review on the use of diffusion models to forecast adoption rate and ownership level of new products see Meade and Islam (2006).

2.1. Gompertz model

Researchers can use a Gompertz model (Gregg, Hossell, & Richardson, 1964) to forecast the car ownership adoption rate. Car ownership level at time t is defined as:

$$N_t = M \cdot \exp(-\alpha \cdot \exp(-\beta \cdot t)) \quad (1)$$

where N_t is the number of cars per 1000 inhabitants, M is the saturation or equilibrium car ownership level in long term where market growth is stagnating and car ownership level reaches a plateau. The two parameters of the model, α and β, define the shape of the growth curve. A defining characteristic of a Gompertz model is that its inflection point occurs before half of the market adopts a new product (Meade & Islam, 1995). This view implies a slow diffusion speed and longer duration period for the diffusion of the new product. Other than the trend variable, which is typically included in the basic specification of the model as in Eq. (1), other explanatory variables can be included. Therefore, Eq. (2) is a general Gompertz model:

$$N_t = M \cdot \exp(-\alpha \cdot \exp(-\beta' \cdot X)) \quad (2)$$