Institution: Cardiff University

Unit of Assessment: 10

Title of case study: Growing Businesses: Robust Models for Understanding Consumer Buying Behaviour

1. Summary of the impact (indicative maximum 100 words)

The School of Mathematics at Cardiff University has developed important statistical and mathematical models for forecasting consumer buying behaviour. Enhancements to classical models, inspired by extensively studying their statistical properties, have allowed us to exploit their vast potential to benefit the sales and marketing strategies of manufacturing and retail organisations. The research has been endorsed and applied by Nielsen, the #1 global market research organisation that provides services to clients in 100 countries. Nielsen has utilised the models to augment profits and retain their globally leading corporate position. This has led to a US$30 million investment and been used to benefit major consumer goods manufacturers such as Pepsi, Kraft, Unilever, Nestlé and Procter & Gamble. Therefore the impact claimed is financial. Moreover, impact is also measurable in terms of public engagement since the work has been disseminated at a wide range of national and international corporate events and conferences. Beneficiaries include Tesco, Sainsbury's, GlaxoSmithKline and Mindshare WW.

2. Underpinning research (indicative maximum 500 words)

The key staff contributing to the research, carried out at Cardiff University during the period 1998-2009, are Professor A. Zhigljavsky (1997-) and Dr V. Savani (PhD student (2002-06), Lecturer (2006-08)). In 2008 Dr Savani was appointed as a full-time permanent statistician at Nielsen.

The basic statistical model we consider is the classical mixed Poisson process with gamma mixing distribution; the so-called Polya process. In the market research context, this model assumes that there are many buyers each purchasing according to a Poisson process with a given rate. The purchasing rates differ from buyer to buyer and are assumed random and mutually independent. For the Polya process, the distribution of the purchase rates is the gamma distribution and the total number of consumer purchase occasions is a negative binomial distribution (NBD). When considering several brands or categories we additionally assume that the buyer chooses the brand according to his or her personal propensities, which are random, independent for different buyers and follow the Dirichlet distribution. Various data sets have been analysed in the published literature demonstrating that this statistical model is appropriate. The immense potential of this model for market research applications had been debated for many years. However, this discussion never progressed beyond a very basic level and as a consequence the model had never before been properly tested and used on data sets of practical value. At Cardiff we extensively studied the statistical properties of this model, in general, with particular emphasis on market research applications for the first time. In particular, our original contributions to the development of statistical models for market research include:

- the development of robust and efficient estimators for the NBD parameters for modelling consumer purchase occasions based on an efficient implementation of the power method; standard methods for estimating the NBD parameters, such as the method of moments and the zero-term method, are inefficient for a significant number of categories. (3.1,3.2);

- the derivation of the joint asymptotic distributions of statistics, including parameter estimators, to study the dynamical behaviour of the mixed Poisson process and hence confidence bounds for the forecasted values of various market research characteristics, including penetration, purchase frequency, mean repeat, repeat per repeater, etc. (3.3); this has required consideration of multivariate distributions with marginal NBD and is essential for assessing the forecasting ability of the basic model;

- the development of various statistical tests for assessing the validity of the model (3.3); testing the Polya process model against integer autoregression and other models with NBD as the marginal distribution (3.2,3.4);

- the development of variations of the basic mixed Poisson model allowing us to deal with seasonal effects, random non-response and the household flow-through problems (3.5, 3.6).

In the course of developing the statistical methodology we tested the methods and models on the largest set of household panel data ever analysed, courtesy of Nielsen. The basic model, statistical tests, estimation methods and forecasting procedures have been modified and adapted in relation to their performance on real data. It was shown that the basic model of mixed Poisson processes often requires certain adjustments to combat problems associated with the presence of non-buyers, the non-stationarity of markets and the panel flow-through. We explored the range of validity of the mixed Poisson model and variations thereof and provided evidence that the model is inadequate at the level of an individual consumer and for small time intervals. However, when sufficient households are aggregated and a reasonable time interval such as a four week period is considered, either the basic model or a variation of it often becomes adequate.

Our statistical findings outlined above have been published in statistical journals (3.1-3.4). Furthermore, in (3.5) we discussed the current state-of-the-art in mathematical modelling in market research, in general, and consumer buying behaviour, in particular. Results of our extensive tests on Nielsen household panel data and adaptations of the model to violations of the main assumptions are described in a series of technical reports submitted to Procter & Gamble (P&G) and Nielsen, for example (3.6). Research funding has come from both Nielsen and Procter & Gamble (3.7).

3. References to the research (indicative maximum of six references)

3.1 Savani, V, Zhigljavsky, A.A. (2006) Efficient estimation of parameters of the negative binomial distribution. Communications in Statistics: Theory and Methods v. 35, No. 4-6, 767--783. http://dx.doi.org/10.1080/03610920500501346 Copy held by HEI, available on request.
3.2 Savani, V., Zhigljavsky, A.A. (2007) Efficient parameter estimation for independent and INAR(1) negative binomial samples. Metrika, v. 65, No.2, 207-255. http://dx.doi.org/10.1007/s00184-006-0071-x Copy held by HEI, available on request.
3.3 Savani V. and Zhigljavsky A. (2007) Asymptotic distributions of statistics and parameter estimates for mixed Poisson processes, Journal of Statistical Planning and Inference, v. 137, No. 12, 3990-4002. http://dx.doi.org/10.1016/j.jspi.2007.04.016 Copy held by HEI, available on request.
3.4 Leonenko N., Savani, V. and Zhigljavsky, A.A. (2007) Autoregressive negative binomial processes, Annales de l'Institut de Statistique de Université de Paris, v. 51, No. 1-2, 25-47. http://cardiff.ac.uk/maths/resources/ISUP_NBDPROCESSES_LeonenkoSavaniZhigljavsky.pdf Copy held by HEI, available on request.
3.5 Zhigljavsky A. (2011) Statistical Modelling in Market Research. In: International Encyclopedia of Statistical Science, Springer, 1450-1452. http://dx.doi.org/10.1007/978-3-642-04898-2_548 Copy held by HEI, available on request.
3.6 Savani, V, Zhigljavsky, A.A. (2007) Confidential Technical Report (ACNielsen BASES): Analyzing Consumer Purchases using Mixed Poisson Models. Copy held by HEI, available on request.
3.7 Research grants with **A Zhigljavsky** as Principal Investigator from

- ACNielsen-BASES on Statistical Modelling in Consumer Studies; 2005, 2006-08, 2008-09, 4 grants totalling US$100k.
- Procter & Gamble on Statistical Modelling in Marketing Research; 1998-2006, 5 grants totalling US$290k.

4. Details of the impact (indicative maximum 750 words)

Route to Impact:

Cardiff University agreed to undertake funded research on sales forecasting for Procter & Gamble, as a result of a long standing relationship, from 1998 and 2006. The outcomes of this research were published in a series of confidential technical reports for Procter & Gamble. In 2004, through the transfer of the Associate Director of Procter and Gamble, CMK (Phil Parker) to Nielsen, the company approached Cardiff University to further develop statistical and mathematical models for forecasting purposes. The research has been applied by Nielsen and has led to the following impacts:

Economic Gain:

Nielsen operates in Africa, Asia, Australia, Europe, Middle East, North America and South America. From 2011-2013, the company has consistently ranked first in the Honomichl Top 25 largest market research organisations in the World. In the group of top 4 global market research organisations, Nielsen research revenue dominates (US$5.4 billion in 2013). Moreover it has increased its global market share by 4% since 2011 outperforming all others in the group (5.1). The products offered by Nielsen (Consumer information, Consumer research and Market measurement) enable companies to understand consumers and consumer behaviour, subsequently increasing profits and expanding businesses. The research conducted at Cardiff University, under an exclusivity agreement, has been integral to the development of Nielsen's services.

Predicting consumer buying behaviour, particularly regarding new or relaunched products, is key to the activities of the corporation. This involves the development of commercially viable analysis and forecasting systems. This is managed by a division called AC Nielson BASES. Prior to 2008 Nielsen was unable to justify a large investment based on the unproven NBD-Dirichlet theory since even the basic mixed Poisson model had not been validated on large household panel data sets. Nielsen required stronger confirmation of the applicability of mixed Poisson models to their data. They commissioned work at Cardiff University by Professor Zhigljavsky and his team to perform an in-depth analysis of the very large data sets provided by Nielsen. The database contains records of every transaction, through the scanning of individual items over a three year period with the variables including date of purchase, quantity, price per unit, flavour, brand and size. The database contains information on over 100 categories and subcategories of products ranging from cereals to soup, cosmetics to detergents and cough treatments to paper products. The statistical analysis of these very large data sets demonstrated that measured behaviour conformed to model predictions and therefore confirmed that the NBD-Dirichlet theory was a viable foundation for effective forecasting and market research analysis. Additionally (see section three) the research provided valuable insights into the most current and efficient mathematical modelling techniques in the market research industry and particularly in consumer buying behaviour.

Nielsen has used the Cardiff research on forecasting consumer buying behaviour to provide services to a host of major corporations. These include Unilever, Procter & Gamble, Coca Cola, Pepsi, Kraft and Nestlé. Nielsen are unable to give explicit information on what services and methods are provided for what clients for reasons of commercial sensitivity. However, the research has been foundational to the work conducted for these global organisations and has enabled the consolidation and progression of corporate relationships. Subsequently, this has enabled Nielsen to compete on a worldwide scale and retain its dominant market position (5.2).

Related research by Professor Zhigljavsky and his team has had significant impact for other companies, such as Procter & Gamble. The models were applied to study sales data versus pricing in several European countries. The models developed at Cardiff University provided more reliable and accurate forecasts compared with the models that were previously employed by the company. For example, the model was used to show that in Italy, sales of a particular Garnier product in drug stores were reduced by modestly increasing the price, whereas in other Italian

outlets increasing the price had minimal effect on sales. In the Netherlands sales of the same product were resilient to modest increases in price. Importantly, this has enabled Procter & Gamble to selectively increase prices in different outlets and different countries without reducing sales. The financial gain associated with this is sizeable.

As a direct consequence of this research Nielsen has invested US$10 million and a further US$20 million in research and forecasting operations. The Vice President of AC Nielsen Product Development (Phil Parker) stated that the research Cardiff has provided has helped them to better understand and address their clients' needs and develop their businesses. He commented that "The positive results of this work have confirmed the theory as a viable foundation for effective forecasting and related analyses, as well as providing insights on efficient methods for analysis and estimation." The subsequent economic rewards that have and will continue to be reaped from this extensive investment have remained confidential but are expected to significantly exceed US$50 million (5.2).

Public/Industry Engagement:

The research has been disseminated by Professor Zhigljavsky at over twenty national and international events, enhancing the knowledge and ability of leading organisations in the manufacturing and retail industries to implement more effective forecasting strategies and respond to a rapidly changing economic climate. These include presentations on the models in Russia, France, Austria, London and Glasgow. Audience members ranged from 30-100 people and included representatives from MindShare WW, MediaCom, DDB Matrix, Ninah, LMG, Billetts / Ebiquity, BrandScience, MacroAnalytica Limited, OHAL and D2D. Moreover, talks have been given to senior staff at GlaxoSmithKline, Sainsbury's, Tesco and the Cooperative (5.3).

5. Sources to corroborate the impact (indicative maximum of 10 references)

5.1 http://www.marketingpower.com/AboutAMA/Pages/2013_GlobalHonomichl25.pdf
Evidence that Nielsen dominates the Market Research Industry.

5.2 Vice President of AC Nielsen Product Development. *Corroborates the use and impact of the research at Nielsen.*

5.3 Corporate Recruiter, Director Briefing Dinner Event (featuring a talk by Professor Zhigljavsky). Pdf document. *Provides an example of public talks given as a consequence of the research*.