

Drawing a Double Jeopardy Line

Cullen Habel and Cam Rungie

This technical note describes the process for drawing a double jeopardy line as it would be estimated by the Dirichlet model. The resultant graph is an x-y plot of penetration and purchase frequency for brands of different market share under a given set of category conditions. Previous analyses show double jeopardy to be an approximately linear effect but the extrapolation inherent in this approach demonstrates the nonlinearity of the effect as brands approach high levels of penetration.

Keywords: Double Jeopardy, Loyalty, Penetration, Dirichlet, Market Modelling

Introduction

One of the most widely recognised empirical generalisations in marketing is that, in Fast Moving Consumer Goods (FMCG) categories, brands or differing market share vary more in the number of customers they have than in any other measure (Bass 1993; Ehrenberg 1972). Brands that have higher market share also have a higher penetration of the available market than those brands with lower market share. Other measures of brand performance (e.g. purchase frequency, share of category requirements) also tend to vary in line with market share but not to the same degree (Ehrenberg, Uncles & Goodhardt 2004).

The nature of these variations has generated strong debate in the marketing literature in the last decade or so. Some authors have taken the market share-penetration correlation to mean that "...a brand's most important customer is the next one it adds" (Anschuetz 2002). Other authors (Baldinger, Blair & Echambadi 2002) recognise that growth of low share brands is predominantly associated with penetration increases, whilst growth of higher share brands tends to be more associated with loyalty increase. The arguments from both sides are convincing and empirically well substantiated. What is lacking from these arguments, however, is a comprehensive understanding of the relationship between penetration and loyalty under a broad range of conditions.

This technical note presents a procedure for describing a theoretical relationship between penetration and loyalty measures for brands within a product category. It makes use of well established, empirically validated models of market behaviour (Goodhardt, Ehrenberg & Chatfield 1984; Jeuland, Bass & Wright 1980). The reader will gain the ability to "draw double jeopardy lines" that apply to their brand's market. The process is applied to a period of antidepressant sales as a demonstration of the process.

This paper considers variables of interest to be:

- Penetration: the percentage of available shoppers who purchase your brand at least once in a given period of observation.
- Average Purchase Frequency: The number of units of your brand that your customers, on average, purchase during the same period of observation.

In the conclusion, we will explain how to extend the process to other variables of interest such as share of category requirements.

Double Jeopardy

Double jeopardy is an empirical result (McPhee 1963). Smaller brands have fewer buyers who purchase less often. The brand with a smaller market share loses twice. It has a smaller penetration and a smaller purchase frequency when compared to brands with bigger market share. Double jeopardy is a basic statistical effect which can be expected to exist in most product categories.

In a graphical manner, double jeopardy can be simply represented as an x-y plot of penetration vs purchase frequency, for brands of differing size. The line of best fit that runs through the scatterplot could be described as the “Double Jeopardy Line”

NBD/BBD and the Dirichlet

This procedure makes use of the penetration/loyalty relationship that is embedded in the NBD/BBD model (Jeuland, Bass & Wright 1980), and the Dirichlet model of buyer behaviour (Goodhardt, Ehrenberg & Chatfield 1984).

Jeuland, Bass and Wright (1980) originally conceived the NBD part of their model with an Erlang interpurchase timing process. This reduces to a Gamma - Poisson (or NBD) as a special case. Both the NBD/BBD and the Dirichlet are empirically well substantiated models of purchase incidence and brand choice that allow the analyst to calculate a range of brand performance measures for brands within a category.

Under these models the Negative Binomial Distribution describes the distribution of category purchase rates across a population of shoppers. The NBD is a two parameter distribution with shape parameter γ (gamma) and scale parameter β . The Dirichlet terminology for these parameters is K and A (Goodhardt, Ehrenberg & Chatfield 1984).

In both models, the brand choice behaviour for the population of shoppers is represented by a multivariate distribution. NBD/BBD represents the two brand case, where the Beta-Binomial Distribution defines the choice probability for brand 1 over brand 2. It uses parameters α_1 and α_2 for the two brands. The Dirichlet represents the multibrand case, with a Dirichlet Multinomial Distribution (DMD) defining the multivariate choice probability for brands within a category. It has parameters $\alpha_1, \alpha_2, \dots, \alpha_n$ for n brands. For a comprehensive statistical summary of the Dirichlet see (Rungie 2003; Rungie & Laurent 2004).

A final parameter of interest is S , which is the sum of the brand's alphas. S has an intuitive application in the Dirichlet in that it is proportional to the rate of switching that occurs between brands in a product category (Sharp & Wright 1999).

The marginal distribution of the DMD is a BBD. Thus if the analysis focuses on the binary choice between one specified brand, say brand 1 and all other brands then the DMD becomes a BBD with parameters α_1 and $S - \alpha_1$ (Johnson, Kotz & Balakrishnan 1997; Kotz, Balakrishnan & Johnson 2000). In this marginal analysis comparing one brand to all other brands in the category the Dirichlet model is equivalent to the NBD/BBD model. This is also shown in the functional forms for penetration and purchase frequency from the Dirichlet model which rely only on the NBD and on the parameters α_1 and $S - \alpha_1$. From the point of view of double jeopardy the Dirichlet and NBD/BBD models are equivalent.

The Relationships Embedded in the Dirichlet and NBD/BBD

Rungie and Goodhardt (2004) defined the algebraic identities embedded in the Dirichlet. These identities act as the “engine room” for the creation of DJ lines. The three that will be used in this procedure are given below:

$$\text{MarketShare}_j = \frac{\alpha_j}{S} \quad \text{Equation 1}$$

$$\text{AveragePurchaseFreq}_j = \frac{\alpha_j \gamma \beta}{S \cdot \text{Penetration}_j} \quad \text{Equation 2}$$

$$\text{Penetration}_j = \sum_{k=1}^{\infty} f_{\gamma, \beta} \left(1 - \frac{\Gamma(S) \Gamma(S - \alpha_j + k)}{\Gamma(S + k) \Gamma(S - \alpha_j)} \right) \quad \text{Equation 3}$$

Where $f_{\gamma, \beta}(k)$ is the probability density function for the category purchase rate, in this case the NBD:

$$f_{\gamma, \beta}(k) = \frac{\Gamma(\gamma + k) \beta^k}{\Gamma(\gamma) k! (1 + \beta)^{(\gamma + k)}} \quad \text{Equation 4}$$

for $k=0, 1, 2, \dots$

How to Draw the Double Jeopardy Line

The procedure for drawing a theoretical double jeopardy line involves holding the parameters γ , β and S constant, allowing the market share of brand 1 to vary from 1 to 99 and calculating the penetration and purchase frequency for each market share value. This change in market share is achieved by varying the α parameter for the brand.

A double jeopardy line is effectively an x-y plot of two variables of interest for brands of different size, at a point in time. We have determined, for the purposes of this paper, those variables of interest to be penetration and purchase frequency.

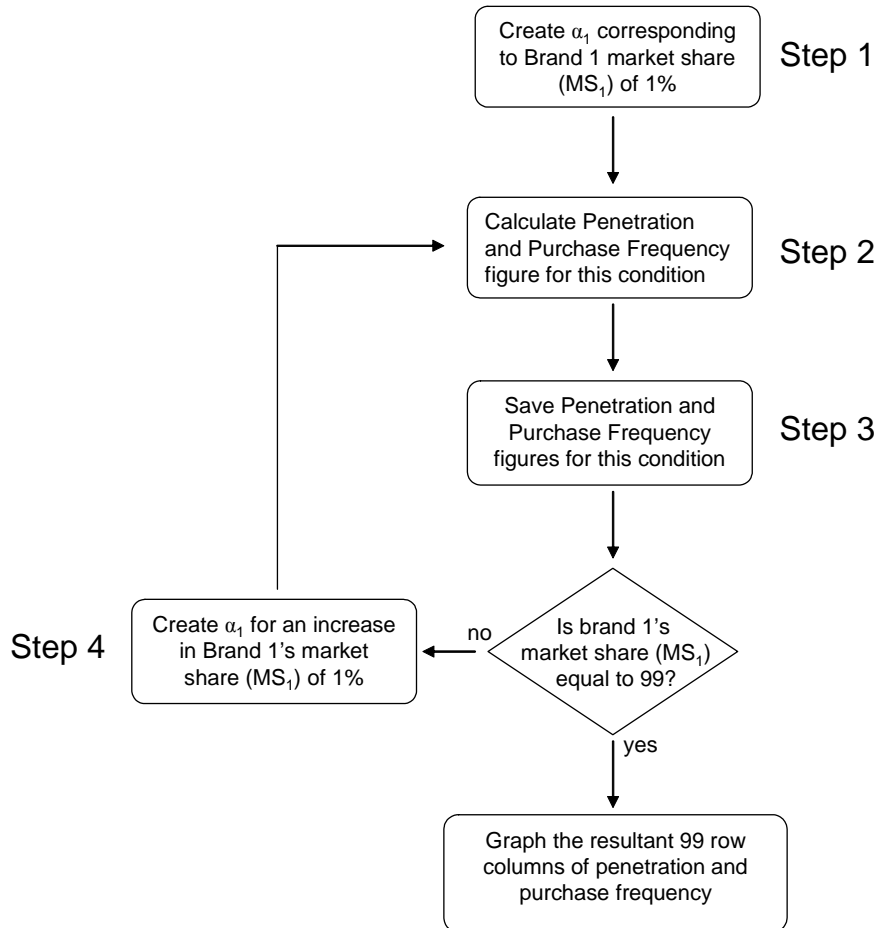
The Inputs to the Procedure

The process of drawing the static double jeopardy line requires three external inputs, being the parameters γ , β and S . These parameters are gaining increasing levels of intuitive meaning (Driesener 2005). A key application of this procedure is to calculate double jeopardy lines to fitted parameters from a dataset, but an analyst with a strong intuitive understanding of typical γ , β and S values may choose to simply arbitrarily specify them. For the purposes of this procedure the parameters will be arbitrarily set as $\gamma=1$, $\beta=1$, $S=2$.

The task can be seen as representing a two brand case, where Brand 1 is allowed to vary its market share from almost 0%, to 100%.

Figure 1 shows the sequence of activities that are required to draw a theoretical DJ line.

Figure 1. The sequence of activities required to draw a theoretical double jeopardy line using NBD/Dirichlet.



Step 1 – Market Share of 1%

Let MS_1 be the value of market share for Brand 1. Recall from the “mathematical relationships” subsection that:

$$MS_1 = \frac{\alpha_1}{S} \quad \text{Equation 5}$$

With a little manipulation, for each value of MS_1 , α_1 can be expressed as:

$$\alpha_1 = S * MS_1 \quad \text{Equation 6}$$

Therefore, with S defined externally as 2 and the market share of brand 1 specified as 1% for step 1, the application of the above equations returns an α_1 figure of 0.005.

Step 2 – Calculate Penetration and Average Purchase Frequency figures.

Using the formulae from the “mathematical relationships” with the externally defined γ , β and S and the α_1 from step 1, the first three values of Market Share, Penetration and Average

Purchase Frequency can be created. For the parameters described above, these three numbers are (.01 .008 1.301).

Step 3 – Save Penetration and Purchase Frequency for this Condition

The row of numbers that is market share, penetration and average purchase frequency that comes about from step 2 is then saved, one row down from the previous entry. In the case of the 1% market share condition, it becomes the top row.

Step 4 – Create α_1 for MS₁ increase of 1%

Using equation 6, α_1 can be calculated for the new market share, being 2% in the case shown above.

Repeat Step 2 and 3

As shown above, the next row of (.02 .015 1.307) can be saved. This process then needs to repeat, increasing market share in increments of 1% and progressively saving the row vectors of market share, penetration and average purchase frequency until a 99x3 matrix is created, as shown in Table 1.

Table 1. Matrix of market share, penetration and average purchase frequency.

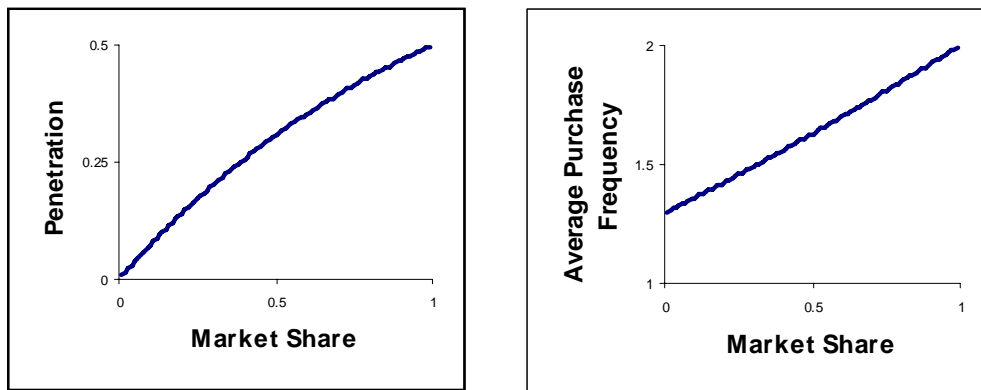
MS	PEN	PF
0.01	0.008	1.301
0.02	0.015	1.307
0.03	0.023	1.313
0.04	0.030	1.320
0.05	0.038	1.326
0.06	0.045	1.333
0.07	0.052	1.339
.		
..		
...		
0.98	0.494	1.985
0.99	0.497	1.992

Step 5 – Graph the resultant 99 row columns

Graphing the figures from Table 1 is simply a matter of presenting the results in the best possible manner. Using the Microsoft Excel graphing function, a number of plots of the response have been created.

Figure 2 shows how both the penetration and average purchase frequency vary in a positive manner with market share increase, albeit with differing slopes.

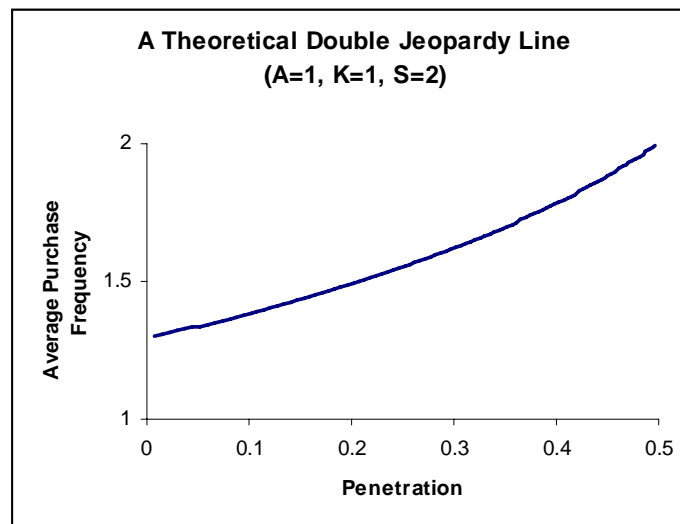
Figure 2. Penetration and Average Purchase Frequency vary positively with Market Share



The fact that penetration increases at a faster (although declining) rate with market share translates to a curvilinear function when average purchase frequency is plotted against penetration.

The final step of the procedure is where the two variables of interest (in this case penetration and average purchase frequency) are plotted together, giving the theoretical double jeopardy line. Market Share varies in the background (see Figure 3).

Figure 3. A Theoretical Double Jeopardy line.



Application

The procedure outlined above can be used to generate theoretical double jeopardy lines for actual markets. What is required to do this is an estimate of β , γ and S for the markets that you are interested in. Recall that β and γ are often referred to as A and K in Dirichlet modelling. There are various means to obtaining these parameter estimates, including the BUYER software (Uncles 1989), the Excel based “Dirichlet” software (Kearns 2002) or likelihood maximisation (Rungie 2003).

We applied this procedure to a three month period in a dataset that documented the prescribing habits of physicians in the United Kingdom during the late 1980's and early 1990's. An extract of the observed and theoretical brand performance measures of interest is given in Table 2. This is the conventional manner of representing double jeopardy. The (O) column is the column of observed values for the given brand performance measure, where (T) is the column of theoretical values, estimated by the Dirichlet model.

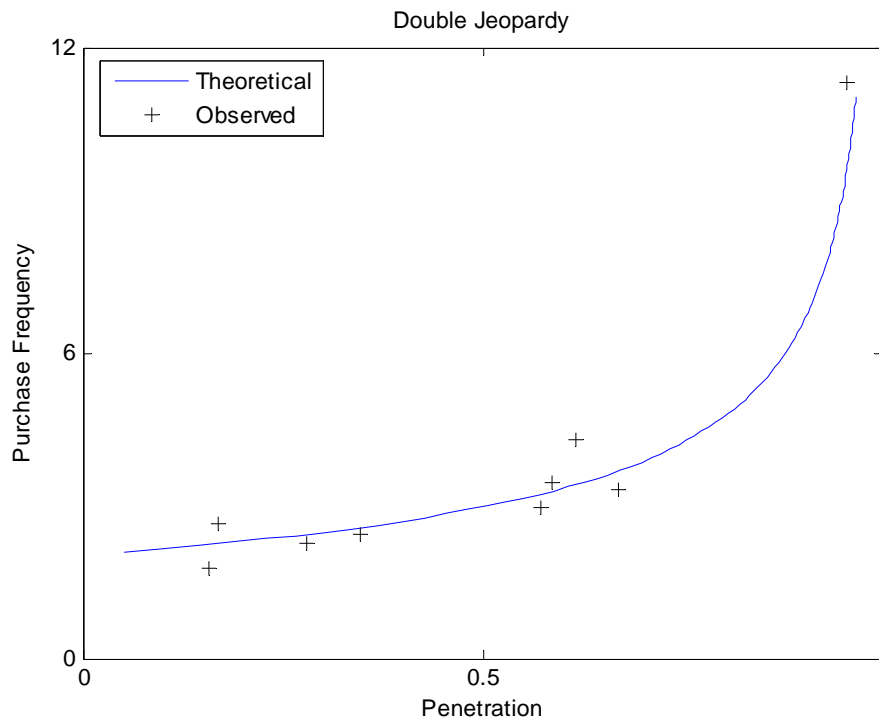
Table 2 shows that penetration varies closely with market share and that there is also a smaller, but real, double jeopardy pattern where the average purchase frequency appears to decline slightly as market share declines. The range of the data is such that the variation in penetration is much greater than the variation in purchase frequency and the double jeopardy effect is, essentially, linear.

Table 2. Theoretical and observed brand performance measures.

BRAND	Market Share (%)	Penetration (%)		Average Purchase Frequency (units)	
	(O)	(O)	(T)	(O)	(T)
B1	25	62	66	4.3	3.7
B7	21	67	65	3.3	3.6
B14	19	59	60	3.5	3.4
B0	16	57	55	3.0	3.2
B4	8	35	34	2.4	2.6
B15	6	28	28	2.3	2.4
B6	4	17	17	2.6	2.3
B18	3	16	15	1.8	2.2
Average	13	42	43	2.9	2.9
CATEGORY	100	95	97	11.3	11.1

The data that is represented in Table 2 can also be presented in a graphical format, using the procedure that is outlined in this paper. Figure 4 shows this. The "observed" data point at approximately 90% penetration is the category penetration and purchase frequency (essentially a brand with 100% market share). The linear double jeopardy pattern that is commonly recognised appears in the range of the observed data, where no brand has greater than a 70% penetration.

Figure 4. A Double Jeopardy line, fitted to Period 30 of antidepressants.



Introducing “doublej.m”

The procedure outlined in Figure 1 has been automated in a matlab executable file that has been named *doublej.m*. It requires an initial specification of the Dirichlet parameters $A(\beta)$, $K(\gamma)$, and S , when the program is called. It then runs a loop that varies market share in increments of 1%, calculates penetration and purchase frequency for each increment returns a 99×3 matrix where each 99 row column represents market share, penetration, and purchase frequency as the brand’s market share goes from 1% to 99%. It makes extensive use of the *dirichletbpm.m* program (Rungie & Goodhardt 2004) to calculate the theoretical brand performance measures. Finally, it plots the three relationships that are given in Figure 2 and Figure 3.

For users of Matlab there are two files attached to this paper. They are “*doublej.m*” and “*dirichletbpm.m*” these need to be copied into the “work” directory of Matlab on the user’s hard drive. The command to run the routine is *doublej(S,[K A])* for S , K and A values that the user chooses to specify. Thus to run a fully automated version of the process that is outlined in Figure 1 the command would be *doublej(2,[1 1])*.

Conclusion

This paper has presented a method for drawing theoretical double jeopardy lines for a given set of Dirichlet or NBD/BBD parameters. The process can be used by an analyst to investigate the penetration-loyalty relationship in a market where those parameters are

known, and make some inferences as to what changes in penetration and purchase frequency may be expected for a given increase (or decrease) in market share. The process has been applied to a market as a demonstration.

It is possible to draw double jeopardy lines for brand performance measures other than penetration and purchase frequency. For instance Allsop (2003) makes a convincing argument for using relative penetration (penetration as a percentage of total category penetration) and share of category requirements. It is a relatively straightforward task to apply other entities contained in Rungie and Goodhardt (2004) when creating Table 1, and thus draw double jeopardy lines reflecting other brand performance measures.

References

- Allsop J (2003). Masters Thesis: Demystifying Double Jeopardy. Marketing Science Centre. University of South Australia.
- Anschuetz N (2002). Why a brand's most valuable customer is the next one it adds. *Journal of Advertising Research*, 42, 15-22.
- Baldinger AL; Blair E & Echambadi R (2002). Why brands grow. *Journal of Advertising Research*, 42, 7.
- Bass FM (1993). The future of research in marketing: Marketing Science. *Journal of Marketing Research*, 30, 1-6.
- Driesener C (2005). Regularities In Repeat Purchase Parameters: Doctoral Thesis. Marketing Science Centre. University of South Australia.
- Ehrenberg ASC (1972). *Repeat Buying* Amsterdam: North Holland.
- Ehrenberg ASC; Uncles MD & Goodhardt GJ (2004). Understanding brand performance measures: using Dirichlet benchmarks. *Journal of Business Research*, 57, 1307-1325.
- Goodhardt GJ; Ehrenberg ASC & Chatfield C (1984). The Dirichlet: A comprehensive model of buying behaviour. *Journal of the Royal Statistical Society*, 147, 621-655.
- Jeuland AP; Bass FM & Wright GP (1980). A multibrand stochastic model compounding heterogeneous Erlang timing and multinomial choice processes. *Operations Research*, 28, 255-277.
- Johnson NL; Kotz S & Balakrishnan N (1997). *Discrete Multivariate Distributions*. New York: John Wiley and Sons.
- Kearns Z (2002). Dirichlet.XLS. Palmerston North: Massey University
- Kotz S; Balakrishnan N & Johnson NL (2000). *Continuous Multivariate Distributions* (2). New York: John Wiley and Sons.

McPhee WN (1963). *Formal Theories of Mass Behaviour* New York: The Free Press of Glencoe.

Rungie CM (2003). How to estimate the parameters of the Dirichlet Model using Likelihood Theory in Excel. *Marketing Bulletin, 14,*

Rungie CM & Goodhardt GJ (2004). Calculation of theoretical brand performance measures from the parameters of the Dirichlet Model. *Marketing Bulletin, 15,*

Rungie CM & Laurent G (2004). Research Note: Statistical Summary of the Generalized Dirichlet Model. *Working Paper, UniSA and Groupe HEC,*

Sharp B & Wright M (1999). There are two types of repeat purchase markets. *28th European Marketing Academy Conference: Berlin, Germany.*

Uncles MD (1989). *Buyer Behaviour Software: A Manual for Version 89.1* London: London Business School, Centre for Marketing and Communication.

Cullen Habel is a Lecturer, and Cam Rungie a Senior Lecturer, in the School of Marketing, University of South Australia