

Bughin, Jacques

Working Paper

Finding the path(s) towards profitable e-commerce

Research Notes, No. 7

Provided in Cooperation with:

Deutsche Bank Research, Frankfurt am Main

Suggested Citation: Bughin, Jacques (2003) : Finding the path(s) towards profitable e-commerce, Research Notes, No. 7, Deutsche Bank Research, Frankfurt a. M.

This Version is available at:

<http://hdl.handle.net/10419/40255>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



January 16, 2003

No. 7

Research Notes

Working Paper Series

Finding the Path(s) towards Profitable E-Commerce

- After the "dot.com" crisis of 2000, some analysts argued, that **e-commerce will only hardly become profitable**. But regarding the performance of business-to-consumer e-commerce (B2C), the current **statistical analyses** demonstrate that many **B2C-companies** have **already achieved** their **turnaround**.
- The empirical analysis shows that **customer conversion and retention** - not online traffic per se - are **the most crucial key factors** for lucrative B2C-commerce.
- Based on a combination of factor and cluster analyses to identify **patterns of behaviour**, three strategies are obvious:
 1. On one extreme, there is a "**dominant reach**" posture, whereby the company focuses on having the **maximum reach** in its segment.
 2. On the other extreme, there is a "**high-value niche**" posture, whereby the company keeps the **focus only on high margin products and customers**.
 3. In-between, there is a "**stuck-in-the-middle**" posture, whereby the **companies attempt to convert traffic into sales at low costs**.
- Although **both extreme postures** - "dominant reach" and "high-value niche" - are usually very **lucrative**, they are only **chosen by a small number** of B2C-companies. In contrast, the "**stuck-in-the-middle**" posture is very **common**, but **only profitable if** - and only if - the **B2C-company has achieved a very high cost efficiency**.

Dr. Jacques Bughin,
Free University of Brussels
Partner, McKinsey & Company



Advisory Committee

Dr. Peter Cornelius,
World Economic Forum

Prof. Soumitra Dutta,
INSEAD

Prof. Dr. Michael Frenkel,
WHU Koblenz

Prof. Dr. Helmut Reisen,
OECD Development Centre

Prof. Dr. Norbert Walter,
Deutsche Bank Research

Deutsche Bank Research

Frankfurt am Main
Germany

Internet: www.dbresearch.com

E-mail: marketing.dbr@db.com

Fax: +49 69 910-31877

Managing Director

Norbert Walter

Finding the Path(s) towards Profitable E-Commerce

Jacques Bughin

*Free University of Brussels, ECARES, Avenue Roosevelt, B-1050 Brussels, Belgium and McKinsey & Company, Avenue Louise, 480, B-Brussels, Belgium **

Abstract

After the shake-out of the years 200-2001, many business pundits have argued - and as much as they have adored it before March 11th, 2000 -, that e-commerce may after all never become a very profitable industry. Relying on a rich and proprietary database e.g., used in **Agrawal et al. (2001)**, this paper takes a closer look at this claim, focusing on the post-entry performance of one category of B2C e-commerce, e-tailers (e-commerce sites that sell transaction goods on-line).

The statistical analyses demonstrate that: (1) Contrary to new wisdom, close to 20% of companies have already achieved operational break-even by mid year 2001; (2): Based on a combination of factor and cluster analyses to identify distinct patterns of competitive behavior within e-tailing, three strategic postures are distinguishable among e-tailers, and in rough consistency with **Porter (1985)**'s theory of clusters. On one extreme, a "*dominant reach*" posture prevails whereby the e-tailer focuses on having the maximum reach in its segment; on the other extreme, one observes a "*high-value niche*" strategy whereby the e-tailer keeps a laser-focus on high margin products and/or customer segments. In-between, the other e-tailers attempt their best to convert traffic into customers at lower possible costs (3): Each of the two extreme postures is chosen by a small core of about 20% of e-tailers, while both postures usually are the most successful in reaching break-even. In contrast, the "stuck-in-the-middle" posture is only profitable provided the e-tailer has achieved strong *cost efficiency* versus average practices. (4) Independently of strategic posture taken, superior conversion and retention capabilities of on-line traffic, *and not traffic per se*, are crucial factors to ensure the post-entry performance of e-tailers.

Keywords: E-tailing, Strategic Clusters, Post entry performance.

* E-mail address: Jacques_Bughin@Mckinsey.com.
The author wants to thank Michael Zeisser, Byron Auguste, Driek Desmet, Marc Singer and Jürgen Wunram for their support.

1. INTRODUCTION

More than three billions of dollars of shareholder value have evaporated from the Nasdaq, since its March 10th, 2000 peak. Among the mostly hurt have been the e-tailers, i.e., retail companies that use the net as a transaction medium. Currently, their stocks just trade slightly above their book value in the US leading many to assert that “*e-tailing is an Icarus myth*” [[Henig (2001)], or [Bruno (2001)]. Not only the financial bubble burst, but even the most supportive of the e-commerce wave, i.e., the Venture Capitalists, have cut their financing to e-tailers by more than two-third their 2000 investments for the subsequent year of 2001.

Yet, on the other hand, some e-tailers have seen their stock price soar back lately by more than 100% such as deliAs, Lastminutes.com, or 1-800-flowers.com [Business Week, (2001)]. Furthermore, the current average life of e-commerce stocks is less than 5 years worldwide [Bughin and Hagel, (2000)], - a too short duration to ascertain whether a company will or will not reach financial viability [see e.g., Mahajin et al., (2001)]. Finally, and as noted by many (e.g. [Klepper and Graddy (1990), or Audretsch (1995)], industries tend to develop within a certain pattern, from a period of explosion to a period of sharp consolidation/concentration that defines industry equilibrium and the set of post-entry winners. E-commerce, - in particular e-tailing -, should be no exception, as after large entry growth rates until May 2000, the trend has drastically reversed with many companies going bankrupt or exiting the market place because of serious lack of profitability, and exhaustion of fundingⁱ.

In order to better gauge the industry pattern for e-tailing, this paper reports on the findings of a statistical research conducted within the frame of the McKinsey DEL initiativeⁱⁱ, on the factors amenable to e-tailing profitability. Leveraging a unique database already used in Agrawal et al. (2001), the various analyses of this paper confirm that operating profitability can indeed be elusive for the average e-tailer, which was still generating a negative operating margin of minus 28 percent of revenue, by mid 2001. Also, the bottom 50% of currently unprofitable companies have generated very strongly negative returns of -110%, while critical operational drivers such as visitor costs, and conversion to on-line buying, have not improved through times, suggesting that those companies will soon need to exit the market. However, also, already a core of 20% of the companies sampled is quite profitable, with margins close to off-line retailing of 20% EBIT (operating earnings before taxes) margin [see Barsh, (2001)].

Concerning the drivers of this profitability, many causes may be found. As is well documented in the business press, the very first reason for gloomy profitability has been that many on-line transaction models simply did not make senseⁱⁱⁱ [see Mahajin, V, R. Srinivazan, and J.Wind (2001)]. The second reason is also that companies have operated significantly under the efficiency frontier; consider for instance the cases of Boo.com, in the UK, or US-born Living.com, who both closed after burning more than 100 million USD on online-marketing campaigns while achieving just the industry average visitor growth or customer conversion rates [Agrawal et al. (2001)].^{iv} Yet, more interestingly, our analysis highlights *strong heterogeneity* of competitive behavior among e-tailers. Consistent with Porter (1985, 2001)'s theory of strategic clusters, (or still with the game-theoretical oligopoly analysis in Mills and Smith (1996) that firms may like to have heterogenous conduct), e-tailers in our sample seem to choose between *three generic postures*. On one extreme, the first strategic posture corresponds to a “*high-value niche*” strategy whereby the e-tailer sells high-margin,

high-tickets, products to a satisfactory core of on-line visitors, while on the other extreme, the third strategic posture is composed of e-tailers choosing a “*dominant reach*” strategy whereby the e-tailer focuses on dominating on-line visitor reach in its segment. In between, we also find a large number of companies, which, while “stuck in the middle”, attempt to follow an “*efficient conversion*” strategy, i.e., they simply attempt to increase the conversion of traffic to on-line buyers, at most possible low costs.

Not surprisingly, each of those postures has achieved its own pay-offs. While the two extreme clusters represent only 20% of e-tailing companies, each in aggregate entails the largest proportion of profitable firms, and especially within the “high-value niche” cluster. On the other hand, the second cluster of “stuck-in-the-middle” companies is only profitable provided significantly lower than average visitor and customer servicing costs. The “dominant reach” is only profitable for those companies, which have converted this reach in a strong customer franchise. In *any* case, superior visitor conversion capability to becoming and retaining an on-line customer is a common operational feature of profitable sites across all strategic postures observed; hence, while many official companies such as Jupiter/MMXI or Nielsen usually track traffic development on-line, conversion, rather than, traffic generation *per se*, remains a must towards profitable e-tailing commerce [Smith et al., (2001), and Moe and Fader, (2001)].

The paper reads as follows. Section 2 presents the sample used in the analysis. Section 3 reports on cluster analysis that leads to the definition of two separate reach and niche strategies in e-tailing. Section 4 investigates the operational drivers to post-entry performance for each cluster. Section 5 concludes.

2. DATA SAMPLE

The present paper leverages a unique (and quite representative) database to measure the post-entry performance of e-tailers along the total funnel of customer attraction and retention. The rough data clearly emphasize that a non-trivial fringe of e-tailers is already profitable, focusing more on traffic conversion to customers and retention than traffic *per se*. However, the variance and skewness in the data suggest large heterogeneity in on-line competitive behavior, which is formally tested and confirmed by cluster analysis in the third section of this paper.

2.1. DATA SOURCES

One usual complaint for a complete economic analysis of on-line business models has been the lack of very good data about Internet companies. On one hand, public-quoted companies publish profit and loss statements; yet nothing is known about key underlying effectiveness drivers of those financial data, such as visitor conversion to customers, repeat visitors, etc... [see e.g. Moe and Fader (2001) or Hagel and Singer (1999)]. On the other hand, if some operational data are ever available (say, from authoritative measurement companies such as Jupiter/ MMXI or Nielsen), they are usually not harmonized, and most of the time only track variations in *traffic aspects* of Internet site performance, such as site unique visitors, page views, or time on site. But again, the foundation for sustained economic returns in e-tailing should be linked to customer lifetime value, which requires to

know a lot more in terms of the dimensions of a company effectiveness, such as traffic-to-customer conversion and customer retention for example.

In order to remedy those caveats, and based on privileged relationships with many companies, McKinsey & Company has launched an '*E-performance*' line since early year 1999, which precisely aims to combine financial data with a full list of operational indicators which track how on-line companies build, maintain and grow effectively their customer base and revenue stream (see **Table 1** for summary indicators collected). The service line has started with B2C sites, and has now extended to B2B sites. As the system is aimed at McKinsey & Company's clients, the individual data and companies included in the full exercise have to be treated as confidential. However, the database includes key industry names, such as for e-tailing, and in random, *eToys*, *JCPenney*, *TheGap.com*, *Marriott.com* in the US, or *3Suisses*, *ChateauOnline*, *Let'sBuyIt* or *BOL* in Europe. Also, it has been claimed elsewhere [**Bughin and Hagel, 2000**, and **Agrawal et al. (2001)**] that the breadth of the survey is sufficiently large to be, if not fully representative, a very illustrative, sample of the Internet economy. In fact, in the cumulative months of January-Sept 2000, the US e-tailing firms in the sample were reported to generate \$8.4 billion dollars of revenue, which is about one-third of total on-line sales as reported by the National Retail Federation/Forrester Research Survey for the same period [**Gertner and Stillman, (2001)**].

Finally, data collection has been done with extreme caution, with up-front harmonized data definition as well as on-site data collection at each participating company, under a consistent time frame (both monthly and quarterly since 1999). The analyses presented in this paper will be based on a sample of about 121 Internet retailing businesses, and for simplicity, most of the Tables included in this article will represent snapshots as of Q1/2001, i.e., for the most frequent data available in the E-performance sample. The sample is distributed globally across North America (45%), Europe (35%), Asia (10%), and Latin America (10%). The e-tailers in the sample are rather young, with less than 4 years of existence on average as of mid- year 2001. Apparels for instance have just 4.1 years, in accordance with **Gertner and Stillman (2001)**, who show that only 4 on the 30 companies in their sample (e.g., Eddie Bauer/Land's End, LL.Bean and J.C. Penney) started operating on-line before 1997^v.

2.2. E-TAILING PROFITABILITY

Table 1 presents some summary statistics for the sampled e-tailers, after splitting the sample between non-profitable (**Table 1.a**) and profitable sites (**Table 1.b**), and where profitability is defined as a on-line site which has achieved positive operating profit (EBIT).

From this **Table 1**, it is visible that the e-tailing sample is still under negative EBIT with average EBIT margin of minus 28% of revenue by mid-year 2001. However, one can notice that the pool of profitable companies in the sample is **N= 23** for **N= 98** for non-profitable e-tailing companies, i.e., $23/121 = 19\%$ of companies are already EBIT profitable. This is on itself a remarkable number, as is possibly larger than the success rate observed by other industries in the beginning of their life cycle [**Klepper and Grady, (1990)**]. Furthermore, the EBIT margin for the average profitable companies is as high as 23 %, which compares very favorably to the 20-25% margins observed in the off-line world for retailers, catalogues and bookstores on average [**Smith et al., 2001**, or **Bughin et al., 2001**].

The flip side is of course that non-profitable e-tailers must have very large negative operating profit margins, i.e. they indeed culminate at minus 38% of revenue. - And even worse, while not reported in **Table 1.a**, the bottom 50% of the non-profitable e-tailers in the e-performance sample reaches negative operating margins of minus 110%. This, combined with a very low, and declining through times, conversion of on-line visitors to customers of just above 1%, makes those bottom-ranked e-tailers unlikely to be profitable and feeds the doubt by many pundits, including the stock market, on the economic sustainability of the e-tailing companies^{vi}.

Table 1 - E-tailing Summary Statistics ^{vii}:

Table 1.a. Non-Profitable Companies

Sub-Sample Size : N=98

Variables	Mean	Std Dev	Minimum	Maximum
Unique Visitors*	373.71	659.11	0.33	6,572.67
Visits per visitors	2.11	0.60	1.2	2.9
Time by Visitors **	16.04	4.02	8.20	24.05
Customers*	7.69	11.24	0.11	61.67
Customers conversion***	2.11			
Customer churn***	77.8	14.3	55.1	89.1
Repeat Customer churn***	24.3	6.5	14.3	54.1
Transactions per customer	0.7	0.5	0.1	2.1
Revenue from transactions ****	1,151.19	4,878.11	42.11	38,545.33
Revenue from others ****	23.22	116.1	0.02	788.32
Total revenue ****	1,185.08			
Total Revenue per customer*****	154.10			
Costs of good sold****	365.01	848.01	0.00	5,951.82
Maintenance Costs****	453.14	574.08	1.45	3,784.33
Marketing Costs****	373.26	593.89	0.08	3,892.67
Other Costs ****	450.08	577.51	5.07	3,784.32
Total Costs****	1,644.50			
Total Costs per customer*****	213.45			
Total Costs per visitor*****	4.39			
Operating margin***	-38.51			

Table 1.b. Profitable companies

Sub-sample size : N = 23

Variable	Profit/ Non- Profit ^{viii}	Mean	Std Dev	Minimum	Maximum
Unique Visitors*	No	255.02	327.53	0.82	5,059.67
Visits per visitors	No	2.95	1.12	2.32	5.12
Time by Visitors**	No	10.40	10.01	2.31	31.23
Customers*	No	27.34	48.63	0.11	145.67
Customers conversion*****	Yes	10.72			
Customers churn***	Yes	53.81	14.56	36.30	70.10
Repeat Customer churn***	Yes	17.12	3.4	11.98	27.89
Transactions per customer	No	1.0	0.5	0.4	2.5
Revenue from Advertising****	No	0.06	0.26	0.00	1.11
Revenues from transactions****	No	2,042.23	3,647.33	12.17	13,100.00
Other Revenues ****	Yes	52.87	224.23	0.23	951.43
Total Revenues****	No	2,095.16			
Total Revenue per customer*****	No	76.64			
Costs of good sold****	No	464.78	1,299.06	3.33	5,700.00
Maintenance costs****	No	416.02	1,120.96	2.57	4,714.00
Marketing costs****	Yes	298.22	580.16	0.00	1,833.33
Other costs ****	No	465.82	1,129.04	2.56	4,312.02
Total Costs****	No	1,614.44			
Total Costs per customer *****	Yes	59.05			
Total Costs per visitor *****	Yes	6.33			
Operating margin***	Yes	26.8			

The first column in **Table 1.b** also presents the result of a simple *univariate* mean T-test of the difference between profitable and non-profitable e-tailing sites. If, arguably, there seem to be not many significant differences (i.e., 8 constructs on 21, or just below 40%), this among others, is linked to the weak power of the non-parametric test involved and the additional observation that the distribution of many of the variables does not converge to a normal-like distribution (see the Min, Max and Standard Deviation values in **Tables 1.a and 1.b**). This may pinpoint to a large heterogeneity in competitive and operational behaviors of e-tailing firms, - an assumption which is formally (and successfully) tested in Section 3 of this paper.

Finally, those early tests tend to show that effective conversion/retention as well as strong cost *efficiency* are all positively correlated with break-even in e-tailing, as discussed hereafter:

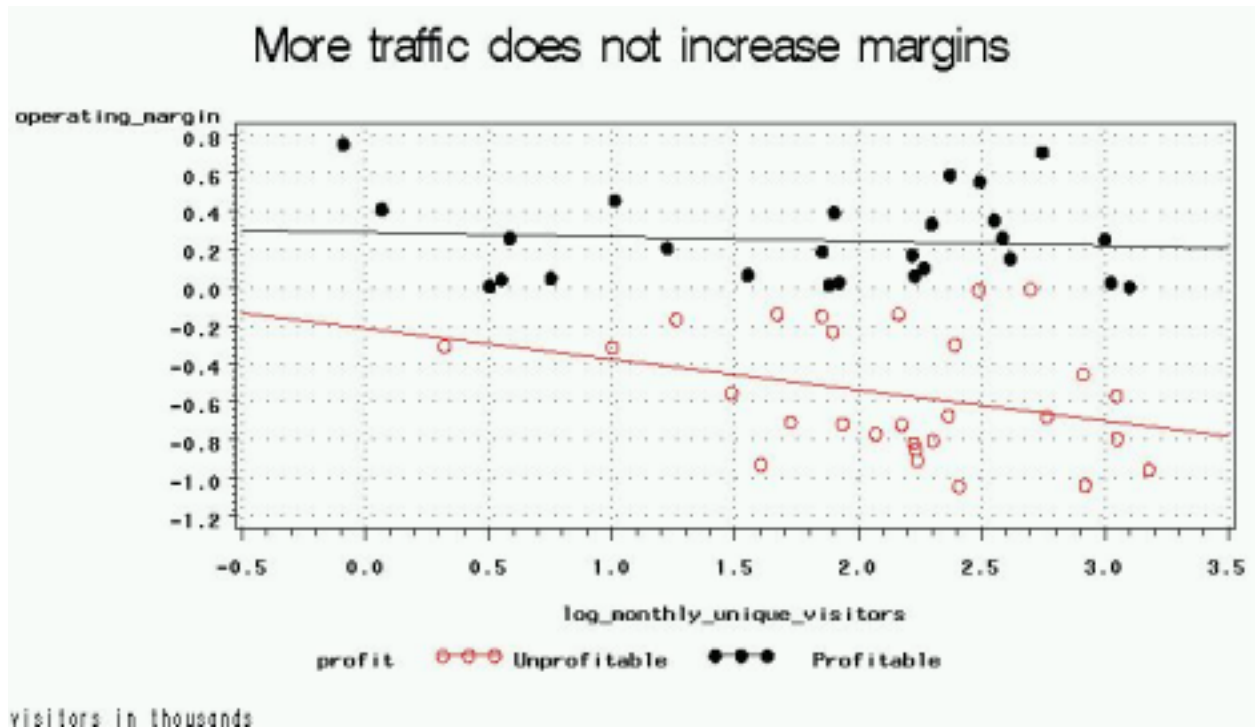
2.2.1. Effectiveness

Traffic Size. The number of average unique visitors is about 315,000 per month. This is typically the average visitors' attraction that one observes for e-tailers in the US according to various data points collectible from companies such as Jupiter/MMXI or Nielsen, but of course, the variance is very large, considering at one extreme sites say, Barnes and Nobles, which currently already attract more than 6 million visits a month. Interestingly, while not statistically different at the mean (see **Table 2**), it remains true that non-profitable sites tend to attract more traffic than profitable ones, i.e., traffic might not necessarily be a sufficient condition for breaking-even. **Bughin and Zeisser (2001)** have looked deeper at this issue and confirm that economies-of-scale in traffic only are rather elusive.

For closer look, **Figure 1** plots operating margin (EBIT divided by revenue) against the monthly unique visitors. A simple regression line based on GMM with firm-fixed effects that predicts operating margin as a function of (log of thousand of) visitors is also reproduced, separately for profitable and unprofitable firms^x. In this present case, the coefficients for the slope are respectively, -0.022 (t-stat=-0.42) and -0.128 (t-stat = -3.04) for profitable and non-profitable sites. Hence, for unprofitable firms, more traffic is indeed significantly linked with much lower operating margin achievement, to the extent that margin is reduced by 12,8 points of percentage when one shifts from a non-profitable site attracting 1,000,000 visitors [i.e., log(3) in thousands] to another one attracting 100,000 visitors a month [i.e., log(2) in thousands]. We will come back to the relationship between traffic and profitability later on in the cluster analysis,

Traffic Stickiness. The average visits per visitors is about 2.5 for about 15 minutes per month, in line with the market estimates such Nielsen and others. Furthermore, profitable companies have larger number of visits per visitor than non-profitable ones (2.95 versus 2.1), yet on-line time stickiness is lower, possibly as a symptom that a larger number of visitors at non-profitable sites are "one-off" visitors, i.e., those who take more time surfing to experiment with the site [see **Lohse et al, (2000)**, or **Bucklin and Sismeiro (2001)**]^x.

Figure 1: On-line Traffic vs. E-tailing Margins



Note: Regression line derived from a GMM regression to predict operating margin in function of traffic, for the full time period.

Customer conversion and retention. More importantly for the economics of retailing is the ability of e-tailers to convert visitors into customers and retain them as an on-going concern franchise. The pairwise T-tests show that customer conversion and retention are statistically different between profitable and non-profitable firms. In fact, concerning conversion, the customer to visitors ratio averages 3.5% for the full aggregate e-tailing sample, in line with other studies [Bellman et al., (1999), Bughin and Hagel, (2000) or Horrigan, (2000)]. However, conversion rate is much higher for profitable sites, at more than 10% versus just above 2% for non-profitable sites, suggesting that customer conversion may indeed be a critical effectiveness factor to reach profitability [Moe and Fader, 2001].

Furthermore, while customer conversion is much larger for profitable sites, churn is also much lower at profitable sites. But, let us notice that in all cases (profitable and non-profitable sites), churn is nevertheless *very large*, suggesting that customer conversion is one thing, customer retention is another more crucial thing. ^{xi}

2.2.2 Efficiency

One final interesting set of figures from **Table 1** relates to efficiency measures. On one hand, revenue *per customer* is not seen to be necessarily higher for profitable than for unprofitable sites. In contrast, however, the cost *per customer* is four times lower for profitable sites than witnessed by unprofitable e-tailers. Combining those efficiency measures, this again confirms that traffic conversion to on-line customer is an important driver for profitability, especially recognizing in **Table 1** that the cost per *visitor* is in fact 40% lower for non-profitable sites than profitable sites.

3. STRATEGIC CLUSTERS

The data above clearly show that effectiveness may draw a clear discriminatory line between profitable and non-profitable sites. However, the sample variance is very large and does not seem to follow a classical normal-like distribution of performance, so that clearly, other elements must be taken into account for assessing profitability. This section investigates the issue of strategic positioning, as developed by **Porter (1985, 2001)** or **Hambrick, (1983)** and more formally described as a strategic answer for differentiation in **Mills and Smith (1996)**.

3.1. EMPIRICAL CLUSTERING

The early view of the internet was that it will create contestable markets, and thus develop low-concentration, cost-minimizing structure. In reality, even e-tailing may not be an efficient market because of consumers heterogeneity in browsing experience for speed, simplicity, information needs, etc. and still a large asymmetry in trust and information. While authors such as **Smith et al. (2001 a,b)** have proven a large dispersion of price in e-tailing proving that e-tailing may not share the feature of a competitive market, **Pan et al. (2001)** also demonstrate that e-tailing companies indeed have developed *differentiation* “look and feel” strategies that lead to different strategic pricing on-line.

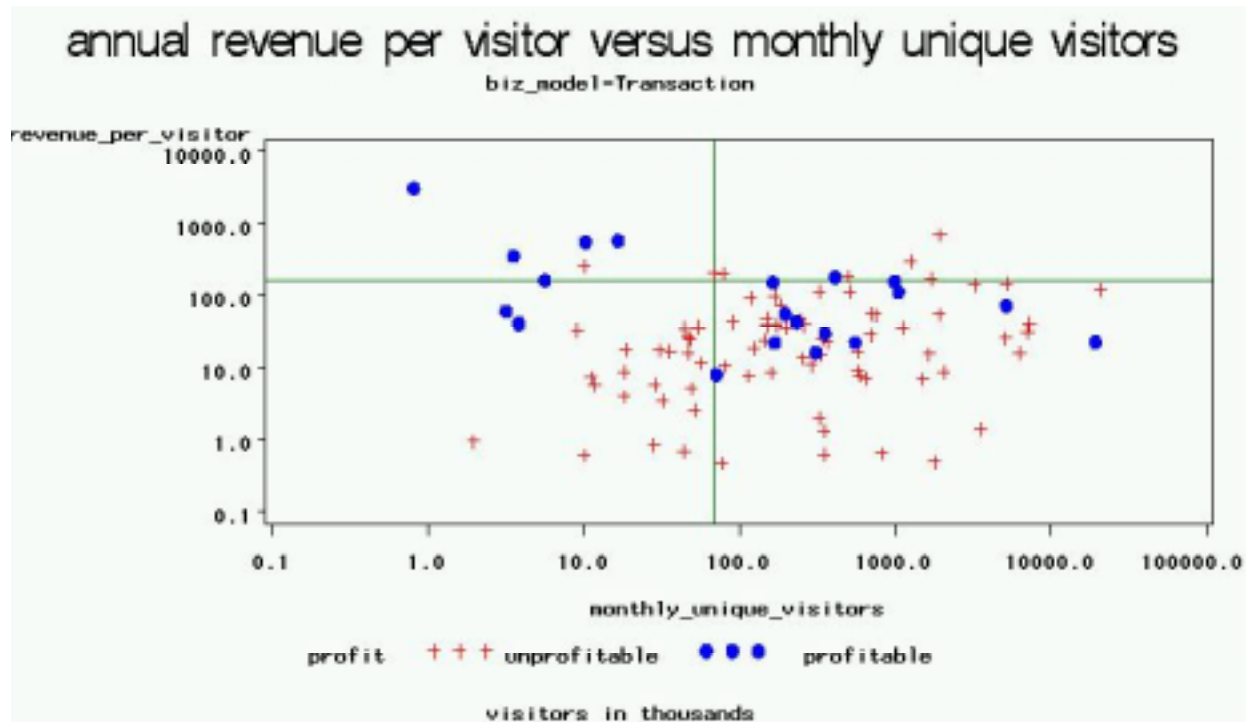
In non-contested markets, **Porter (1985)** argues for three generic business-level strategies that will define competitive advantages. Those are *differentiation* (with the aim to create brand loyalty and price inelasticity) , *cost leadership* (which requires high market share and large efficiency), and *focus* (which aims at segments within an industry, whether a specific customer segment, a product line, etc.). **Rubach and Moore (1998)** confirm that off-line retailers tend to reproduce the generic strategies as in the seminal work by **Porter (1985)**, so we hypothesize that various strategic clusters may also be emerging on-line.

A visualization of our sample can already distinguish at least two separate profitability paths for online transaction retailers: one based on high revenue per visitor (but low unique visitors), and another one, characterized by lower revenue per visitor but large visitor base **[Figure 2]**.

In fact, the visual plot in **Figure 2** indicates that companies which generate more than 100 US dollars annually, (or more than 8 US dollars monthly) are in all likelihood profitable, especially when they do

not attract more than 100,000 unique visitors a month. Since, typically, fixed costs of e-tailing averages about 1.0 million USD of yearly expenses [Agrawal et al, (2001)], this also means that profitable sites below 100,000 visitors have achieved strong conversion capability, at least 10% of visitors converted to cover fixed costs expenses - hence, we hypothesize that those profitable sites should represent niches of quite strong franchise on-line (see infra).

Figure 2 - On line Traffic vs. Revenue per visitor



In the same way, for sites which attract more than 100,000 unique customers, average revenue per visitor may still be cut by two versus low visitors sites and still be profitable - this comes of course directly from a larger potential of dilution of the average fixed costs per visitor described above. However, the plot also means that, especially for sites which have been attracting more than one million unique customers, it is increasingly difficult to generate profitable margin as well as to sustain good revenue per visitor in consistency with **Figure 1** above. Those sites are typically “high reach” sites with three times more visitors than the average, and we conjecture that they should be linked to low-ticket items products given ceteris paribus, lower than average revenue^{xii}.

3.2. STATISTICAL CLUSTERING

To formally identify strategic clusters, we first need to reduce the information contained in all “E-performance” variables. We thus have run factor analyses in order to find the minimal number of dimensions which may explain the maximum amount of the overall covariance of the observed variables. Using a classical Varimax rotation methodology, and constraining the factor loading to be above 0.5 (i.e, a traditional norm), four factors have been retained that explain more than 90% of the data variance. Finally, based on factor loading, the score summary of each e-tailing company has been clustered based on their similarity, and has generated three specific K-means different clusters, in consistency with **Figures 1 and 2** above.

3.3.1. Factor analysis

The factor analysis is reproduced in **Table 2**. For interpretability, all cost items correlations are with respect to the inverse of the cost item, so that a *positive loading stands for cost-efficiency*.

The first factor correlates strongly with all conversion and retention attributes. As such, this factor corresponds to the conversion effectiveness capability of an e-tailer. The second factor corresponds to the dimension of revenue maximization, given how the factor correlates *strongly and positively* to high revenue per customer or per transaction. Clearly, this factor replicates one dimension of **Figure 2** above and should be correlated with stronger profitability (see infra). The third factor is the reflection of the attraction “focus” of the e-tailer, as the factor correlates positively with visitors and reach (reach bring as visitor reach at product segment level). Finally, the last factor correlates with low marketing as well as maintenance cost per visitor (as well as customer). This matches the dimension of “cost leadership”.

3.3.2. Clustering

Table 3 reports the K-means cluster analysis as per 1st quarter of 2001. It emphasizes three *strategic* clusters around the four factorial dimensions above, with the ANOVA results for difference in the mean of each factor dimension, being largely significant across the three clusters found (all have p-values below the threshold, 1/1000) ^{xiii}. The clusters are also very stable through times, and present a good balance of two strong extreme clusters (which are composed of about 20% each of the companies in the E-performance sample) and one large group of companies which is like “stuck in the middle”.

On one extreme, the first of the smallest clusters (*Cluster 1* in **Table 3**) is comprised of e-tailers with a special focus on revenue maximization, and still strong conversion capability, but which are not especially concerned with high traffic. This cluster matches the cluster of companies with very small traffic in **Figure 2**. It is thus called the “*High-value niche*” strategy cluster as it is concerned with a small franchise base for high-ticket selling items only; in fact, more than 45% of our on-line groceries and apparel sites (versus an a priori uniform norm of 20%) are located in this segment together with some luxury goods e-tailing companies.

On the other extreme of the spectrum, *Cluster 3* with a total of 20% of the sampled e-tailers, encompasses all companies which are quite focused on traffic building. This cluster is mostly composed of e-tailers which on average sell low-ticket items (34% of movie/books/CDs companies are located in this segments), but as well, some higher-ticket goods and services (45% of travel sites found in this cluster). We call this cluster, the “*dominant reach*” strategy cluster, as companies are mostly concerned with establishing large traffic reach.

Finally, in between, we find *Cluster 2* which is comprised of all other “*stuck-in-the middle*” e-tailing companies. In general, those companies tend to have slightly better than average cost efficiency, but do not seem particularly effective in generating strong reach, revenue, or even, conversion capabilities.

Table 2: Factor Loadings

	Factor1	Factor2	Factor3	Factor 4
Label	“Conversion ”	“Revenue Capability”	“Traffic Focus”	“Cost leadership”
Dimensions:				
Customers	0.90060	-0.04480	0.18172	-0.12673
Transactions	0.95226	-0.10984	0.12553	-0.14602
Repeat visitors	0.92374	-0.10924	0.12283	0.21965
Repeat customers	0.59990	-0.12103	0.04937	0.12032
Repeat customer conversion	0.52045	0.19319	-0.19824	-0.17432
Revenue per trans.	-0.11427	0.98784	0.09481	0.17894
Revenue per cust	-0.06899	0.85861	0.00484	-0.12567
Unique visitors	0.25385	-0.13093	0.94384	0.14691
Visitor percent reach	-0.13864	-0.16703	0.79371	-0.22912
Marketing cost per visitor*	-0.31781	0.34612	0,15742	0.69421
Cust mktg cost*	-0.11577	-0.32850	0.09079	0.74508
Maintenance cost per visitor*	-0.17650	-0.13451	0.11405	0.81256
Eigenvalues	4.3945143	2.9994544	2.7690465	2.046329

Notes: Only dimensions with loadings above 0.5 for at least one factor shown. Correlation shown for 1rd quarter 2001.

*** Correlations shown are with the inverse of the cost construct.**

3.3.3. Clusters' Profitability

What are the pay-offs for each of the strategic posture identified by the cluster analysis? To this end, **Table 3** also illustrates the average profitability achieved *within* each of the clusters as well as the *proportion* of companies that have reached break-even by the first quarter of 2001.

	Cluster 1	Cluster 2	Cluster 3
"Conversion Capability"	0.47	-0.21	0.10
"Revenue Maximisation"	1.65	-0.66	0.15
"Traffic Focus"	-0.44	0.11	1,05
"Cost Focus"	-0.23	1.12	0,04
Cluster size	18%	62%	20%
Proportion of profitable companies			
Within Cluster*:	41%	15%	18%
In total of companies:	38%	47%	15%
Average operating margin	-6.7%*	-23.7%*	-60.3%*
Of which:			
Profitable Operating margin	27.3%	18,9%	24,2%
Non profitable operating margin	18.9%**	-32.3%**	-74.5%**

Notes: Figures based on 1rst quarter, 2001.

*: T-tests of cluster dummies in logit profitability model are all statistically significant

** : Saitterwhaite mean T-tests at unequal variance are all statistically different

Furthermore, **Table 4** reports the mean of some operational features between non-profitable and profitable companies *per each cluster*.^{xiv}

By far, the "high-value niche" provides the best average operating margin performance with an EBIT which is only negative at 7 points of margin. The "reach" posture has achieved a quite negative EBIT, in line with **Figure 1** above.

Interestingly, the major driver of this average difference among clusters lies mostly in the *proportion* of firms that have already been able to reach break-even, (and is confirmed through a simple logit model that predicts profitability in function of each cluster dummy. See Table A.2 in appendix); in fact, *given* the fact that they have achieved break-even, e-tailing margins converge to 20% of EBIT for all clusters, but the percentage of profitable firms is significantly higher for the "high-value niche".

Looking finally at the results of an analysis of the features that separate profitable from non-profitable e-tailers within in each separate cluster, **Table 4** confirms the existence of three major idiosyncratic paths towards profitability, one for each cluster:

Table 4: Profitable E-tailing Features by Cluster

	Cluster 1	Cluster 2	Cluster 3
Unique Visitors*	17.235	218.640	758.921
Time by visitors**	40.35	16.18	22.71
Customer conversion***	13.19	5.23	8.37
Transactions per customer	1.42	1.10	1.16
Revenue per transaction****	182.12	75.00	79.15
Customers churn***	53.21	53.46	58.92
Visitor marketing costs****	4.82	1.34	0.17
Customer maintenance costs****	27.26	18.21	25.00

Note: Data are averaged out for the three months of 1st quarter 2001. All data are monthly except churn which is quarterly. All dimensions in table 4 are statistically significant drivers of profitability per cluster.

* in thousands; ** hours; *** percentage; **** in US dollars

Path 1: Succeeding with a “high-value” niche strategy. Table 4 shows that compared to non-profitable sites, profitable e-tailers adopting a “high value niche” have typically invested quite high costs of marketing at 5 Euro per monthly visitor, mostly as a result of spreading marketing costs around a relatively small visitor base (about 25,000 unique monthly visitors). However, this high-cost strategy is matched with significant high revenue flows as the profitable “high-value niche” e-tailers manage to extract significant revenue per customer (more than 180 USD dollars per transaction and 1.4 transaction per month per each customer acquired) while also have achieved strong practice of conversion and retention, at average of 13% of their visitors and of 47% of quarterly customers respectively.

Since the conditions above for a profitable “niche” requires significant revenue per transaction and enough buying frequency, groceries (food and beverage) and general apparel, may seem the only categories generating annual spending likely to allow a business to generate 150 USD in gross margin per visitor, and *taking into account double digit conversion rates*. While other categories are unlikely to support profitable “niche strategies” on a stand-alone basis, they might nevertheless sustain profitable companies, if those manage to aggregate many products to attain a critical share of consumer wallet, or still manage to target to specific consumer segments whose annual spending in a category is significantly higher than that of the average population. A case example here can be the pet shop *Waggin’ Tails* [Business Week, 2001] which has specialized in hard to-find, 30% something margin products, including vitamins and high-end dog foods, and is now profitable with \$5 million

sales [compare to revenue achieved by profitable e-tailers in our niche posture] versus the now-shut down *Pets.com*, which has tried an universal, low margin on-line reach strategy delivery [see below].

Path 2: Succeeding within a dominant reach strategy. As *Pets.com* has witnessed versus *Waggin'Tails*, the economics of a “dominant reach” strategy are rather the opposite of the ones of a high-value niche” strategy in the dimensions of traffic attraction and revenue generation. Hence, the subset of e-tailers which manage to be profitable in the “dominant reach” posture have attracted a large number of visitors (more than 750,000 unique monthly visitors), which is however needed to compensate for the more limited revenue (typically 80 USD per transaction).

In consistency, however, with the profitable “high-value niche” companies, the profitable companies in this strategic posture also have achieved quite strong conversion of their traffic, at more than 8% rate, but also have spread costs enough for marketing costs to be a few cents only.

Path 3: Succeeding in the middle segment. Then remain all the “stuck in the middle” e-tailing companies. While in all likelihood struggling to establish a stable franchise or to develop a very large on-line basis of visitors, a set of companies have managed to reach profitability. Those focus especially on compensating for lack of reach (220,000 monthly visitors) or high-value positioning (lower revenue per transaction and lower transaction per customer than other profitable companies in other clusters) by being relatively cost-efficient; visitor acquisition cost is only 17 cents, or less than 5 times the average of the total E-performance sample, while customer maintenance costs are also less than 20 USD dollar per customer.

In general, the three paths share the common feature, they all share above average retention and conversion capability.

5. CONCLUSIONS

Relying on a unique and arguably representative sample of transaction e-commerce companies, this paper has taken a closer look at whether and how companies could be profitable on-line. Our analysis clearly has demonstrated that a non-trivial set of winners are emerging, which combine both strong operational excellence as well as fine-tuned to about 3 generic strategic postures.

The paper also proves that customer conversion and retention are critical effectiveness factors to reach profitable economics, but not traffic per se, on transactional e-commerce.

In subsequent research, we intend to extend the analysis to content sites, while refining the E-performance database on critical drivers of effectiveness dimensions, such as conversion and retention capabilities. The distinction between “dot.coms’ and “dot.corps” will also be exploited.

6. REFERENCES

- Agrawal, V. L.Arjona, and R. Lemmens (2001)**, "E-performance: the path to rational exuberance", The McKinsey Quarterly, 2001 (1).
- Arellano, M. and S. Bond (1991)**, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", Review of Economic Studies, 18.
- Audretsch, D.B, (1995)**, "Innovation, growth and survival", International Journal of Industrial Organization, (13)4, pp. 441-458.
- Barsh, J. (2001)**, "How e-tailing can re- birth from ashes", The McKinsey Quarterly 2001(1).
- Bruno, L. (2001)**, "B2C looks for the right fit", Red Herring, June 12.
- Bucklin, R. and C. Sismeiro, (2001)**, "A model of web site browsing behavior estimated on clickstream data", June, mimeo, Anderson School, UCLA.
- Bughin, J, and J.H. Hagel (2000)**, "The operational effectiveness of virtual communities", International Journal of Electronic Markets, Dec.
- Bughin, J., S; Hasker, L.Segel and M.Zeisser (2001)** ,"Reversing the digital Slide", The McKinsey Quarterly Anthologies,Media and Entertainment: Byte-Sized Strategies, pp. 47-56.
- Bughin, J, and M. Zeisser (2001)**, "The marketing scale effectiveness of virtual communities", International Journal of Electronic Markets, Dec.
- Bughin, J (2002)**, "Can on-line sites be an economic success? Monte Carlo evidence", mimeo, McKinsey Digital Economy Lab
- Business Week (2001)**, "Rethinking the internet – special report", March 26.
- Demers, E. and B. Lev, (2001)**, "A rude awakening: internet shake-out in 2000", Working paper, University of Rochester.
- Gertner, R.W. and R.S. Stillman (2001)**, "Vertical integration and internet strategies in the apparel industry", Journal of Industrial Economics, XLIX, 4, Dec., pp. 417-440
- Hambrick, D.(1983)**, "An empirical typology of mature industrial-product environments", Academy of Management Journal, 26, pp.213-230
- Henig, P. (2000)**, "Is e-tailing dead?", Red Herring, March 20.
- Horrigan, J. (2000)**, "New Internet Users- what they do, don't on-line and implications for the net's future", Pew Internet and American Life Project.
- Klepper, S, and E.Graddy, (1990)**, "The evolution of new industries and the determinants of market structure", Rand Journal of Economics, 21, pp. 27-44.
- Klepper, S, and J.H. Miller, (1995)**, "Entry, exit, and shakeouts in the US in new manufactured products", International Journal of Industrial Organization (13)4, pp. 567-592.
- Lohse, G, Bellman, S. and E. Johnson (2000)**, "Consumer behavior on the internet- findings from panel data", Journal of Interactive Marketing (forthcoming).
- Mahajin, V, R. Srinivazan, and J.Wind (2001)**, "The dot.com retail failures of 2000: were there any winners?" Working Paper, McCombs School of Business, University of Texas at Austin

- Mills, D. and W. Smith (1996)**, "*It pays to be different: endogenous heterogeneity of firms in an oligopoly*", International Journal of Industrial Organization, 14(3), pp. 317-330.
- Moe, W. and P. Fader**, (2001) "*Capturing evolving visit behavior in clickstream data*", The Wharton School, Marketing Department Working Paper, March 2001
- Pan,X., B. Ratchford and V. Shankar (2001)**, "*Why Aren t the Prices of the Same Item the Same at Me.com and You.com?: Drivers of Price Dispersion Among E-Tailers*", mimeo, Sloan School of Management.
- Porter, M. (1985)**, "*Competitive strategy: techniques for analyzing industries and competitors*", New-York:The Free Press.
- Porter, M. (2001)**, "*Strategy and the Internet*", Harvard Business Review, March, pp. 63-78
- Ruback,M. and J.McGee (1998)**, "*Stuck in the middle: for retailers, perhaps not such a bad place to be*", mimeo, university of Texas.
- Smith, M.D, J.Bailey, and E. Brynjolfsson (2001)**, "*Understanding digital markets- Review and assessment*", in "*Understanding the digital economy*", MIT press, Brynjolfsson and Kalin, eds. (forthcoming).
- Smith, M.D, J.Bailey, and E. Brynjolfsson (2001)**, "*Consumer decision making at an internet shopbot: brand still matters*", Journal of Industrial Economics, XLIX,4, Dec., pp. 541-558

APPENDIX

This appendix lays out the supporting statistical analyses referred to in the main body of the paper.

Table A.1: EBIT versus on-line traffic GMM regression

	Profitable sites	Non-Profitable sites
Fime-effects	Yes	Yes
Time-effects	Yes	Yes
Dot.corp dummy**	0.104* (2.46)	0.026* (3.66)
Log of monthly visitors in thousands:	-0.022 (-0.42)	-0.128* (-3.04)
EBIT-1	0.86* (3.66)	0.92* (6.47)
N (0,1)	0.26	0.64
Pseudo -R ²	0.72	0.61

- Estimation based on first-differencing fixed effects, and after instrumenting EBIT-1 with lagged variable t-2; monthly visitors and geographic dummies (Asia, US, Europe)
- N (0,1) is a Normal test for second-order serial correlation, and is statistically non significant.
- T-statistics in brackets; * means significant at $\alpha = 5\%$; ** dummy equals 1,0 otherwise if e-tailer is a “dot.corp” company (pure “dot.com” otherwise).

Table A.2: Break-even logit regression

	Cluster only	Cluster and interaction with operational variables
Cluster 1 dummy	0.36 (45.12)	0.20 (16.33)
Cluster 1 times above revenue per transaction		0.24 (16.56)
Cluster 1 times above transaction per customer		0.06 (6.17)
Cluster 1 times below marketing costs		0.12* (1.31)
Cluster 1 times above customer conversion		0.17 (53.22)
Cluster 1 times above time by visitors		0.09 (14.12)
Cluster 2 dummy	0.12 (7.12)	0.07 (6.91)
Cluster 2 times above revenue per transaction		-0.02* (1.06)
Cluster 2 times above transaction per customer		0.31 (17.14)
Cluster 2 times below marketing costs		0.12* (1.31)
Cluster 2 times above customer conversion		0.14 (12.67)
Cluster 2 times above time by visitors		-0.06* (-1.72)
Cluster 3 dummy	0.20 (14.36)	0.16 (16.26)
Cluster 3 times above revenue per transaction		0.09 (14.12)
Cluster 3 times above transaction per customer		0.04* (0.62)
Cluster 3 times below marketing costs		0.06 (2.21)
Cluster 3 times above customer conversion		0.26 (19.42)
Cluster 3 times above time by visitors		0.03 (3.22)
Likelihood ratio (chi-square)	1,246	2,843
Score (chi-square)	1,422	2,712
Wald (chi-square)	503	1,624
Percent concordant predicted and observed	66,3	84.2

- Estimation based on Fischer scoring with logit function (1= profitable; 0=non-profitable). Convergence criteria satisfied for all equations. Coefficient in brackets are Chi-square values.
- * means not significant at $\alpha = 10\%$

ⁱ Internet bankruptcies more than doubled between 2000 to 2001, according to Webmergers (www.webmergers.com). The year-end survey published by Dec. 2001 declared that more than 530 Internet companies shut down or declared bankruptcy in 2001, more than twice the 225 that shuttered operations in 2000. Interestingly, although the data would suggest a major winnowing of internet companies, Webmergers argues otherwise. By the company's own estimates, between 6,000 Internet companies received some sort of "formal" funding during the same two-year period, so in total, only 15 % shut down in last two years. This may mean that the consolidation is not necessarily finished yet, as historically, the empirical findings converge to a threshold of 35% of firms remaining after shake-out and consolidation [see e.g., **Audretsch (1995)**, or **Klepper and Miller (1995)**].

ⁱⁱ DEL, for "Digital Economy Lab" is an initiative launched by McKinsey & Company in 2001 and which analyses possible winning strategies in the New Economy. Some of its research findings have appeared in various issues of the McKinsey Quarterly, Harvard Business Review or Electronic Markets.

ⁱⁱⁱ Consider the illustrative example of "hotsauce.com", a retail start-up. Their business plan seeks to generate \$500,000 in sales in its first year. If the company realized average transaction value of \$25, it would need to close 20,000 transactions to generate its target \$500,000 in revenue. With a customer to visitor ratio of 1%, hotsauce.com would need to attract two million visitors to reach its first-year revenue goal. If hotsauce.com relied solely on banner advertising with a click-through rate of 0.5%, hotsauce.com need to purchase 400 million impressions to net two million visitors. At an average cost of \$0.01 per banner ad (\$10 cpm), it could cost \$4 million to generate sales of \$500,000. This does not include expenses for things like site development, maintenance and the cost of goods sold. While the example is oversimplified, the point is clear: online retailing does not make sense in certain markets.

^{iv} **Bughin and Zeisser (2001)** report that marketing effectiveness is indeed very poor in B2C, even for virtual communities which should have exploited the viral aspect of marketing spent. They report large univariate diseconomies-of-scale for marketing spent with respect to visitor traffic and revenue generation.

^v Incidentally, there are all companies who did also launch catalog operations before on-line.

^{vi} **Bughin (2002)** uses Monte-Carlo techniques to better grasp the distribution of survival profitability of e-commerce companies. Concerning e-tailers, a company with less than 1.5% conversion ratio has virtually a zero probability of reaching break-even.

^{vii} All data are monthly, averaged out over the three months of the first quarter 2001, except for customer churn, which is quarterly. Customer conversion as well as revenues and costs per customer/visitor are computed at the sample mean.*: In thousands; ** hours; *** Percentage; **** In thousands of USD dollars; *****: In USD dollars.

^{viii} Satterwhaite pairwise t-test. Yes (No) means that the difference between profitable and non-profitable sites is (not) significant at 5% risk level.

^{ix} The semi-log form is the one that maximizes the univariate fit. This has also been statistically checked through former Box-Cox tests. **Table A.1** in appendix presents the full tests.

^x We also have collected page views for the companies in our sample—this tends to be about 8 per visit for profitable sites and only 4 for non-profitable sites.

^{xi} Assume for once that marketing costs are constant per unique targeted customer. Since, on average, marketing costs per customer represents 20% of revenue, a repeat customer with at least twice more transactions than a one-off customer will at least generate 10 points of percentage more of operating margin, outside any other saving from diluting per-customer fixed costs.

^{xii} The fixed-effects GMM estimate of the regression slope that predicts revenue per visitor in function of traffic is -82 with an absolute T-test value of 2.61 (for GMM application to economics see **Arellano and Bond (1991)**). The regression is of the semi-log form to maximize best fit.

^{xiii} The F-tests are for illustrative purposes only –see references in **Pan et al. (2001)**.

^{xiv} Except for unique visitors, most of the variables are significant predictors of break-even within each cluster. This has been checked using a battery of tests, from the simple equal mean T-test, to logit analysis to predict break-even (see **Table A.2** in appendix). Also, a non parametric discriminant analysis on the variables laid out in Table 4 classifies all profitable companies accurately, and only 6% of non-profitable companies were misclassified as profitable for the total E-performance sample.

All our publications can be accessed, free of charge, on our website www.dbresearch.com. You can also register there to receive our publications regularly by e-mail.

© 2003. Publisher: Deutsche Bank AG, DB Research, D-60272 Frankfurt am Main, Federal Republic of Germany, editor and publisher, all rights reserved. When quoting please cite „Deutsche Bank Research“.

The information contained in this publication is derived from carefully selected public sources we believe are reasonable. We do not guarantee its accuracy or completeness, and nothing in this report shall be construed to be a representation of such a guarantee. Any opinions expressed reflect the current judgement of the author, and do not necessarily reflect the opinion of Deutsche Bank AG or any of its subsidiaries and affiliates. The opinions presented are subject to change without notice. Neither Deutsche Bank AG nor its subsidiaries/affiliates accept any responsibility for liabilities arising from use of this document or its contents. Deutsche Banc Alex Brown Inc. has accepted responsibility for the distribution of this report in the United States under applicable requirements. Deutsche Bank AG London being regulated by the Securities and Futures Authority for the content of its investment banking business in the United Kingdom, and being a member of the London Stock Exchange, has, as designated, accepted responsibility for the distribution of this report in the United Kingdom under applicable requirements. Deutsche Bank AG, Sydney branch, has accepted responsibility for the distribution of this report in Australia under applicable requirements.

Printed by: HST Offsetdruck Schadt & Tetzlaff GbR, Dieburg

Print: ISSN 1610-1502 / Elektr.: 1610-1480 / Internet: ISSN 1610-1499