Value in Direct Navigation: Empirical Evidence

Alex Tajirian

October 5, 2006

ABSTRACT
The paper points out the limits of two empirical studies on the value of direct navigation. To more accurately predict the value created through direct navigation, these issues must be addressed.

INTRODUCTION

Direct navigation has tremendous marketing implications[1] and is of paramount interest to buyers and sellers of Internet properties (domain names and websites). Thus, it is imperative that the sources and magnitude of additional value creation are investigated thoroughly.

There are two major empirical studies on direct navigation. The first, by WebSideStory,[2] finds that sales-conversion rates when a visitor arrives at a site through direct navigation are higher than other forms of Internet navigation. The second[3] considers only the relationship between direct navigation and its impact on domain name prices and parking revenue.

WebSideStory was the first to use sales-conversion rates as a proxy for comparative navigational revenue. The proxy is an intelligent choice for website profits, which are not publicly available. The study also developed an interesting visitor classification model. Earlier studies used comparative traffic volumes across various Internet navigational vehicles, which lacked any measures of profits.

Both studies have limitations that are detailed below.

WebSideStory Study

The study classifies visitors into three levels, and estimates the conversion rates for each level and across various product groups. However, the study has a number of shortcomings with regard to visitor classification modeling, estimation techniques, and interpretations of results.

---

[1] See Alex Tajirian, "Direct Navigation: Marketing Implications," DomainMart.

[2] WebSideStory, Search Engines Have More Than Twice the Conversion Rate of Other Acquisition Sources, Press Release (January 30, 2006).

[3] See Alex Tajirian, "Direct Navigation Hubris?," DomainMart.

ClickZ quotes Ali Behnam, WebSideStory senior digital marketing consultant[4], providing a description of these variables. Level two consists of "people who have an idea of what they want, are in the market for a specific product, but don't necessarily know where they want to get it." Level three, on the other hand, consists of "people who know what they're looking for, and know where they want to shop to get it."

1. Classification Modeling

a. From the perspective of modeling visitors, the study's verbal classifications of variables and states can be translated into the 2x2 framework below. Although such models are simplifications, they are a very powerful analytical tool.

Visitor Classification

		Know What They Want	
		No	**Yes**
Know Where to Go	**No**	Level 1 (?)	Level 2
	Yes	Level 4	Level 3

Thus, the framework results in four disjoint visitor levels.

The level 4 group, however, is not mentioned in the study! Compulsive shoppers, for example, fall in this category, whereby they may regularly visit a favorite discount/bargain website.

b. The study only considers same-session conversions, which makes the distinction between direct navigation and search engine transactions blurry. For example, the study cannot tell whether a direct navigation had originated with a search engine visit. In such a case, search engine conversion rate would be biased downward. Conversely, it does not account for the fact that a visitor might have started the search process with direct navigation, but, say, did not find the intended site and decided to use a search engine. Ironically, arriving at a search engine site, on the second leg of the search, would most likely be through direct navigation!

2. Estimation

Conversion rates can be formulated into the following investigative question: What is the conversion ratio for each of the visitor levels?

[4] Direct Navigators Convert Better than Searchers, ClickZ News (January 31, 2006).

To answer the above question, there are two methodological approaches: using a controlled experiment setting and data mining techniques. The latter is the approach adopted by the study.

a. Controlled Experiment Approach

An experiment can be conducted in which visitors' knowledge of where to go and what they want can be explicitly controlled using a focus group. Nevertheless, the limitation of this approach is that while it can control visitor types, unless the focus group's participants make purchases too, estimated sales-conversion rates can be suspicious, as the focus group's items that they intend to purchase and purchasing the items need not correlate.

b. Data Mining Approach

The study is based on tracking activities of about 30 business-to-consumer ecommerce sites over the last three months of 2005.

i. Measurement Error

In a data mining setting, to estimate the sales-conversion rate for each visitor "what they want" and "what they know" are held constant by the statistical[5] techniques. In such an approach, when control variables of the visitor classifications have to be imputed, visitor classification becomes less accurate, and thus, errors in control variables[6] must be included in the statistical test.

For example in the statement that "[level three] users converted almost twice as well as those using search engine links," suggests that level 3 users use both search engines and direct navigation. Thus, how can the researcher ascertain the level of the visitor? The same identification problem exists for level 2 users. Moreover, if someone were truly level 3, (i.e., know where they want to go) why would they use a search engine? One plausible reason is that they are not sure of the exact spelling or the extension of the destination's domain name. Thus, they would use a search engine to find the address of the site that they know they want to go to. They may then bookmark it for future visits. So, in the study's classification scheme, would such a visitor be classified as a direct navigation person or search engine? If the former, it would bias the level 3 conversion numbers in favor of direct navigation.

Hence, for the estimation results to be meaningful, the investigator has to either integrate the measurement error into the estimation procedure or has

[5] In a regression experimental design, for example, the control variables are included in the regression as independent variables.

[6] The issue needs to be statistically addressed. For example, using *ordinary least squares* regressions in the presence of such errors results in biased estimates.

to be certain that a person using a search engine can be classified as either a level one, two, or three (or four) visitor.

ii. Multicolinearity

Depending on the test design formulation, there can be high correlation between the navigation method and visitors' knowledge of where they want to shop online. In a regression setting, for example, such multicolinearity leads to serious estimation reliability problems.

3. Interpretation of Results

a. Product Categories

Visitor behavior for five product categories is tested for differences in sales-convergence rates. The study points out the importance of "familiarity" of the site as a determinant in purchasing a big-ticket item compared to, say, purchasing toys. But under level 3, it is assumed that the visitor knows where they are going. Thus, what does "familiarity" represent for each of the categories and how is it measured? Moreover, should level three visitors, for example, be redefined based on the three characteristics "know what they want," "know where to go," and "familiar with the site"? If so, this familiarity factor is another variable that needs to be imputed, adding to the problem of reliability of estimates.

Nevertheless, the investigation into the reason for variations in conversion rate across groups becomes interesting. One possible explanation is that for bigger ticket items, the searcher bookmarks more sites or visits more sites through type-ins than for small-ticket items. Thus, the conversion rate is lower due to more bookmarks and/or type-ins for the big-ticket items. This explanation makes sense, as the number of sites visited for big-ticket items is probably higher due to potentially significant cost/price-quality variations.

b. Conversion Rates

Concentrating on the conversion rate, while ignoring marketing costs, can be misleading. Some of the type-ins and bookmarks have been paid for by previous marketing campaigns. Thus, when such costs are ignored, even with higher conversion rates, it does not necessarily imply that direct navigation creates additional value to the visited sites.

DomainMart Study

The study suggests that direct navigation does not add value to a parked domain name. The lack of significance of direct navigation can be attributed to statistical estimation issues, as noted in the study, and that publicly available domain name sale prices do not reflect the full benefits of direct navigation, i.e., sellers, and possibly buyers, on average, do not fully understand the value of direct navigation. ■