MARKETING SCIENCE
Vol. 26, No. 6, November–December 2007, pp. 758–773
ISSN 0732-2399 | EISSN 1526-548X | 07 | 2606 | 0758

informs®

DOI 10.1287/mksc.1070.0286

The Value of Quality

Gerard J. Tellis
The Marshall School of Business, MC 0443, University of Southern California, Los Angeles, California 90089-0443, tellis@usc.edu

Joseph Johnson
508 Kosar Epstein Building, Coral Gables, Florida 33146, jjohnson@miami.edu

Product quality is probably undervalued by firms because there is little consensus about appropriate measures and methods to research quality. We suggest that published ratings of a product's quality are a valid source of quality information with important strategic and financial impact. We test this thesis by an event analysis of abnormal returns to stock prices of firms whose new products are evaluated in *The Wall Street Journal*. Quality has a strong immediate effect on abnormal returns, which is substantially higher than that for other marketing events assessed in prior studies. Moreover, there are some important asymmetries in the effect. We discuss the research, managerial, investing, and policy implications.

Key words: stock returns; quality; published reviews; new products
History: This paper was received on December 16, 2004, and was with the authors 12 months for 4 revisions; processed by Scott Neslin.

Introduction

Firms are always in a rush to bring new products to market. This attitude is probably driven by several factors, such as capturing market share rewards to new products; shaping consumer preferences; exploiting economies of experience; or preempting lucrative supplies, market positions, distribution outlets, and shelf space (Robinson and Fornell 1985, Carpenter and Nakamoto 1989, Kalyanaram et al. 1995). However, compelling evidence and some theory suggest that first movers may not have the advantages attributed to them in the past (Golder and Tellis 1993, Shankar 1999, Zhang and Narasimhan 2000). Further research has shown that speed alone does not lead to higher sales growth, nor does it lead to higher accounting returns (Ittner and Larcker 1997). Superior quality is what consumers really look for in a new product (Liebowitz and Margolis 1990, Tellis et al. 2007). Indeed, quality may well be the most important factor in the success of products and the market performance of competing brands (Tellis and Golder 2001). Yet firms may systematically undervalue the importance of quality.

Why is this so? Several reasons may account for why firms may undervalue quality. First, new product introductions are glamorous, and being the first mover is prestigious (Tellis and Golder 2001). Second, quality is difficult to define, measure, and assess unambiguously (Curry and Faulds 1986). Third, market response to quality is not instantaneous but occurs over a long time (Mitra and Golder 2006). Fourth, there is no universal agreement on what quality means. Indeed, quality is studied differently in several academic paradigms, including engineering excellence, perceived quality (Dodds et al. 1991), reviewed quality (Tellis and Wernerfelt 1987), self-reported quality (Tellis and Fornell 1988), consumer satisfaction (Rust et al. 1999), and brand loyalty (Aaker 1995).

Thus, some highly pertinent questions have remained unanswered: Is there a reasonable summary measure of quality? Are there underlying product dimensions that explain such a summary measure? Does the market recognize and value such a measure of quality? Is this value fully realized immediately or over a period of time?

No research has yet addressed these questions, primarily because of the difficulty of measuring quality. In contrast, a number of studies in marketing have assessed stock market returns to new product introductions, new product preannouncements, use of celebrity endorsers, use of Internet channel, brand extensions, firm name change, and joint ventures and contracts (see Table 1).

This paper suggests a simple but powerful method to measure quality and assess market rewards to it: the event analysis of abnormal stock market returns of firms whose new products are publicly and systematically rated. Use of stock prices for market valuation has at least three benefits. First, data on stock prices are abundant and precise. Second, an accepted paradigm of research provides a clear method—the event study—for assessing stock market returns to

Table 1 Event Studies in Marketing

Authors	Topic of study	Sample size	Effect size on day of event (%)
Horsky and Swyngedouw (1987)	Company name change	58	0.61
Chaney et al. (1991)	New product introductions	1685	0.72
Agarwal and Kamakura (1995)	Celebrity endorsements	110	0.54
Lane and Jacobson (1995)	Brand extensions	89	0.18
Houston and Johnson (2000)	Joint ventures and contracts	583	0.02
Geyskens et al. (2002)	Introduction of Internet channel	98	0.35

information about quality. Third, a focus on stock prices is responsive to the ultimate goal of the firm—maximizing shareholder value (Srivastava et al. 1998).

The event study is so called because it attempts to evaluate the impact of an event on the returns of a stock. Typically, an event represents the arrival of new information to the market. Though event studies have been around for a long time (Dolley 1933), Ball and Brown (1968), and especially Fama et al. (1969) laid the economic foundation for the method in use today. Since then, event studies have been widely used in finance to study the impact of different classes of informational events on the returns of stocks. Fama (1991, p. 1607) says that the "cleanest evidence on market-efficiency comes from event studies, especially on daily returns." The use of the method has spread from finance to accounting, organizational behavior, strategy, and other fields.

There are limited but increasing applications of this method in marketing (see Table 1). In addition to these studies, Pauwels et al. (2004) use time series methods to relate new product introductions to firm value. However, they do not assess how quality of products affects stock market value. Indeed, no published study has examined stock market returns to product quality using the event study. This paper reports on such a study.

The next six sections explain the theory, method, data, results, implications, and limitations of the study.

Theory: Event Study of Reviewed Quality

This section briefly explains the logic of the event study. It then develops hypotheses of the impact on the stock prices of firms when information about the quality of their products hits the market.

Events, Information, and Market Efficiency

An efficient market is defined as one in which the market price is an unbiased estimate of true value. Implicit in this definition is the assumption that prices reflect all information available to investors. Under the most widely accepted interpretation of market efficiency, known as the semistrong version, market prices reflect all available public information about a stock, including both past and current information (Fama 1970, 1998). This concept of efficiency implies that market prices can only change because of the arrival of new or unanticipated information affecting the future prospects of a stock. It also implies that the speed with which new information is incorporated into prices is instantaneous and that the magnitude of the price change is a measure of the value of this information. Figure 1 displays the market reaction to new information as envisioned in the theory of efficient markets.

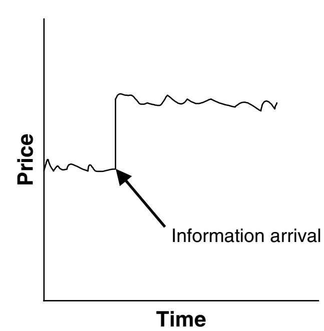
This average movement of stock prices to the arrival of information is called *normal returns*. The returns of any individual stock that differs from those of the market are called *abnormal returns*. This study focuses on how information about the quality of the new products of a firm affects its abnormal returns. The following subsections develop rival hypotheses about our expectations in this regard.

Abnormal Returns and Reviewed Quality

Our expectations about how information on quality affects stock prices are defined by two rival theories. We call these theories heterogeneity of tastes and imperfect information.

Heterogeneity of Tastes, Reviewed Quality, and Returns. There is unanimity in marketing that quality is a multidimensional construct (e.g., Curry and Faulds 1986, Kopalle and Hoffman 1992). Many

Figure 1 Reactions to Unanticipated Information Arrivals in an Efficient Market



authors strongly assert that consumers differ on their preferences for the various dimensions (Hjorth-Anderson 1984, Kamakura et al. 1988). Such authors assume that no single expert can possibly rate the quality of products unambiguously because he or she would not be able to come up with a composite scale that would appeal to all consumers. If that were the case, then a measure of composite quality would have no particular relationship to any underlying dimensions for quality. Thus, there are strong arguments for the following null hypothesis:

Hypothesis H_{01}. *A composite measure of quality obtained from public reviews of quality will bear no relationship to and cannot be composed from underlying quality dimensions.*

If the above is true, then the stock market would not be influenced by any published report on the ratings or reviews of composite quality, however good the expert writing that report. Thus, ratings or reviews of composite quality would not lead to any abnormal returns to the stocks of the parent firms, on the day they were released to the public. Thus, there are strong arguments in favor of the following null hypothesis:

Hypothesis H_{02}. *There are no abnormal returns to firms whose composite quality is publicly reviewed.*

However, other researchers have argued strongly that the concept of composite quality is valid and relevant for two reasons. First, for a composite measure of quality to be invalid, both of the underlying conditions have to be met (Kopalle and Hoffman 1992): (i) The dimensions have to be strongly negatively correlated, and (ii) the weights or utility on those dimensions by consumers or raters has to be negatively correlated. Second, such conditions are very rare in real markets (Curry and Faulds 1986). Thus, one might posit the following rival Hypothesis to H_{01}:

Hypothesis H_{A1}. *A composite measure of quality from published reviews of quality will be positively related to its underlying dimensions.*

Imperfect Information, Reviewed Quality, and Returns. Investing in quality to gain a competitive advantage is a primary concern of management (Mittal et al. 2005). Research conducted in economics, marketing, and management indicates that superior quality leads to high performance (Shaked and Sutton 1983, Sutton 1986, Metrick and Zeckhauser 1998). Specifically, studies in marketing show that composite quality exerts a significant positive influence on market share (Jacobson and Aaker 1987, Buzzell and Wiersema 1981, Phillips et al. 1983, Kordupleski et al. 1993), return on investment (Buzzell et al. 1975, Phillips et al. 1983), and price (Phillips et al. 1983, Tellis and Wernerfelt 1987).

At the same time, ample evidence suggests that consumers are imperfectly informed of quality and cannot assess quality immediately by inspection (Nelson 1970, Tellis and Wernerfelt 1987). In such cases, experts with knowledge and resources could sample products before consumers do and evaluate them through personal experience or formal experiments. Such an exercise would endow the experts with knowledge about the quality of those products that is superior to that of the average consumer. If these experts published their evaluations of composite quality, then uninformed consumers could rely on them. The extensive availability of and demand for reviews and ratings of electronic products in computer magazines and on the Internet and of all products in *Consumer Reports* is testament to the market value of such reviews (Mayzlin 2006, Moorman et al. 2005, Eliashberg and Shugan 1997).

If the above argument is valid *and if* the market is efficient, then investors would understand the importance of reviews of quality. Thus, when information on the quality of a firm's product becomes available in the market, the price of that firm's stock is likely to go up or down depending on the review. If the information is favorable, we would expect investors to warm up to the stock of the firm, thus raising its value on the date of the rating's release. On the contrary, if the information is unfavorable, we would expect investors to pull out of their positions in these firms, causing a drop in their values. Thus we would expect abnormal returns of that stock depending on information about its product composite. So, our alternate hypothesis is as follows:

Hypothesis H_{A2}. *The abnormal returns to a firm's stock whose composite quality is publicly reviewed vary with the rating in the review.*

However, if experts are unable to correctly value quality or if consumers do not care for such information, then the null Hypothesis H_{01} would hold.

Negativity Bias

Would abnormal returns vary for negative versus positive reviews of quality? A growing literature asserts that consumers react more negatively to losses than positively to gains (Kahneman and Tversky 1979). This hypothesis has been called the negativity bias. If this hypothesis carries through to the context of abnormal returns, it would be reflected in steeper negative abnormal returns to poor reviews of quality than positive returns to good reviews of quality. The effect may carry through because either investors themselves react this way or investors foresee that consumers would react this way. Thus, a testable hypothesis is:

Hypothesis H_{A3}. *The abnormal returns for inferior reviews of quality are more negative and greater in absolute value than those for superior reviews of quality.*

Reputational Asymmetry

Investors react to news that surprises them. Thus, prior expectations affect how new information impacts stock prices. People normally tend to interpret new information in a manner that is consistent with or supports their prior expectations (Einhorn and Hogarth 1978). Because of their order of entry or prior performance, small firms probably receive lower appreciation and loyalty than large firms (Moorman et al. 2005). Prior expectations about small firms are likely to be more negative than those for large firms. In fact, fewer investors track small firms than large firms (Shefrin 2000).

When good news about the product of a small firm reaches the market, investors are likely to be positively surprised, causing prices of such firms to rise. In contrast, when investors receive negative information about the product of a large firm, they are likely to be negatively surprised, thus pulling down stock prices of such large firms. These opposite reactions would result in asymmetric response to the reviews of quality of small and large firms. Thus, we hypothesize:

Hypothesis H_{A4}. *Inferior reviews of quality lead to more negative abnormal returns for large firms than for small firms; conversely, superior reviews of quality lead to more positive abnormal returns for small firms than for large firms.*

Relative Importance of Dimensions of Quality

Research on the relative importance of various dimensions of quality in marketing is nonexistent. Thus, hypotheses about such issues are tentative. In general, for technology-based products, the most important dimensions of quality would be compatibility, ease of use, and utility of features, for the following reasons. Firms are constantly coming out with new generations of hardware and software products. However, consumers rarely replace all their hardware and software at once. Rather, they replace them as their needs change or as a substantially new product emerges. At that point in time, backward compatibility with older products is of paramount importance to consumers. Second, new products tend to be developed by techies with techies in mind; they tend to be rather complex and not user friendly. Thus, user friendliness would be a very important dimension of quality for consumers. Third, new high-tech products also come out with a bewildering variety of bells and whistles. Some of these are very useful, others trivial, others confusing. Thus, utility of features would be very important for consumers. In contrast, because high-tech products do not have to last very long (typically warranties cover most of their useful lives) and because they are generally more stable than mechanical products, we do not expect reliability to be an important dimension (unlike, for example, for cars or washing machines).

Method

This section first describes the context and the event for the study. It then describes how abnormal returns of a stock (beyond what are normal for the market) are calculated.

Context

Firms typically do not systematically release information about the quality of their products. Thus, the first problem in assessing abnormal returns to quality is obtaining a systematic and standard format for the market release of information on quality of products, at least for new products.

One solution to this problem is to assess the stock market returns to published reviews. Fortunately, new electronic products are frequently introduced on the basis of superior performance on several objective criteria. Because of consumers' need for information on these products, several publications compete with each other to regularly and carefully rate new products. If the market is efficient, then the abnormal returns to information on quality should reflect the true underlying worth of that quality (e.g., Aaker and Jacobson 1994, Fama 1970, Malkiel 1992).

For more than a decade, *The Wall Street Journal* has published an insightful review by Walter Mossberg of the quality of electronic products and software. The review has been consistent in that it occurred every Thursday, had the same length, and was placed on the first column of §B. The long duration and consistency of the latter reviews suggest that consumers value them for reflecting underlying quality. The placement of these reviews in *The Wall Street Journal* suggests that investors would access this information. Thus, the reviews could affect the price of the parent stocks.

Here are two quotes that partially attest to some of these premises:

> Walt Mossberg is walking through a convention hall at the Consumer Electronics Show in Las Vegas when a man starts screaming at him. The screamer, Hugh Panero, blames Mossberg for his firm's recent problems: falling stock price, a sudden plunge in consumer interest. Mossberg is annoyed but hardly intimidated. As the author of the weekly "Personal Technology" column in *The Wall Street Journal*, he's used to dealing with disgruntled execs. He lets Panero shout. A crowd is gathering. Finally, Mossberg yells back, "I don't give a **** about your stock price!" (Deutschman 2004, p. 1).

> If the consumer technology business has a guiding patriarch it is Walter Mossberg. His "Personal Technology" column in the *The Wall Street Journal* is without doubt the most influential and widely read source of commentary on new technology products. No other technology journalist has Mossberg's clout or his power to influence technology design itself; for example, Microsoft dropped Smart Tags after Mossberg attacked it as an abuse of power.... Mossberg gained this kind of sway through a combination of approachable writing, brutal honesty, and unquestioned objectivity—all characteristics I have to admire deeply as a fellow technology writer. (Roush 2004, p. 1).

The context of our study is new electronic products and software, whose quality was reviewed by Walter Mossberg regularly in the Thursday issues of *The Wall Street Journal* from 1991 to 2001. We assess the abnormal returns, if any, to the stocks of the parent firms caused by these reviews. Note that if this event did not represent substantially new information, then it should not trigger any significant abnormal returns. Indeed, the great noise in the market loads the test in favor of the null hypothesis of not finding any effects.

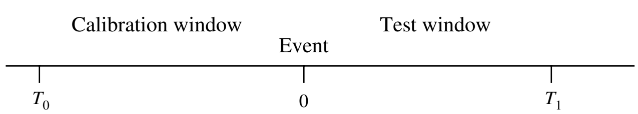
Measuring the Event

All our analyses are at the daily level. Such a disaggregate level of analysis increases the power of the test and reduces the potential for biases. Note that adopting daily analysis does not prevent or compromise the analysis of long-term effects. Figure 2 graphically portrays the definitions of the event and related periods for our study. The event is the day of the publication in *The Wall Street Journal* of the review of quality of a new product by Walter Mossberg. We count that event as day 0. In all cases, prices are the closing ones for any day.

The calibration window is the period over which we estimate how a stock normally relates to the market as information reaches the market (Equation (3)). The test window is the period over which we estimate the abnormal returns, if any, caused by the event. These abnormal returns are computed as deviations between the actual returns of a stock and the predicted returns based on its relationship to the market estimated in the calibration period (Equation (4) below).

In our particular study, to capture any leakage before or decay after the event, we used a window starting from five days before the event to five days after it. A window larger than five days runs the risk of contamination with information from the next week's events. For the calibration period, we used a 180-day window prior to the 6 days before the event. Technically, the larger this calibration window, the more stable the estimates.

Figure 2 Events and Windows



Model: Derivation and Testing of Abnormal Returns

We define returns per unit time, R_{it}, measured at time period t, for a firm, i, as a change in price P between $t-1$ and t, plus dividends at time t D_{it}, and adjusting for splits. Thus

$$R_{it} = (P_{it} - P_{it-1} + D_{it})/P_{it-1}. \tag{1}$$

P_{it} is the price that is adjusted for stock splits. This computation is done by multiplying the price P_{it} by an adjustment factor (facpr). This factor computes the number of additional shares per old share issued. It is calculated as follows:

$$\text{facpr} = \frac{s_t - s_{t-1}}{s_{t-1}} = \frac{s_t}{s_{t-1} - 1}, \tag{2}$$

where

s_t = number of shares after split
s_{t-1} = number of shares before split.

We try to capture what portion of this return is *"normal"* or common to all stocks in the market. Classical financial theory holds that the return to a firm's stock is determined by the unique performance of each stock and the performance of the general market from economic, political, and environmental conditions; thus:

$$R_{it} = \alpha_i + \beta_i * R_{mt} + \varepsilon_{it}, \tag{3}$$

where

R_{mt} = market rate of return, i.e., the average return of the New York Stock Exchange (NYSE) or Standard & Poor's 500 at month t
α_i = the time invariant idiosyncratic effect of firm i on its own return
β_i = effect of the entire market on the return of firm i
ε_{it} = errors with $\mathrm{E}[\varepsilon_{it}] = 0$ and $\mathrm{Var}[\varepsilon_{it}] = \sigma_\varepsilon^2$
t = -180 to -6 (for calibration period).

Now, the predicted value of Equation (3) is the "normal" component of a stock's return, or the portion that is driven by the unique performance of each stock plus the market as a whole. The difference between actual returns and projected returns from Equation (3), if any, gives the *abnormal returns* (AR_{it}). Thus:

$$\mathrm{AR}_{it} = R_{it} - E[R_{it}], \tag{4}$$

where

AR_{it} = abnormal return for stock i on day t
R_{it} = actual return for stock i on day t

$E[R_{it}]$ = expected return for stock i on day t predicted by Equation (3)
$t = -5$ days to $+5$ days.

Abnormal return measures the impact of information release over a given day and is thus a poor indicator of the total impact of the information release. Cumulative abnormal return, which is the sum of all abnormal returns over the time period of interest, is a better indicator of the full impact of the information release. We calculate the cumulative abnormal return ($\mathrm{CAR}_{i\tau}$) by cumulating the abnormal returns for firm i over a period of time τ, thus

$$\mathrm{CAR}_{i\tau} = \sum_{t=0}^{\tau} \mathrm{AR}_{it}. \qquad (5a)$$

If the abnormal returns of a portfolio of stocks are statistically related to any event that is common to all stocks in that portfolio, we can conclude that an event was the cause of the abnormal returns. To test this hypothesis, we pool the cumulative abnormal returns over the different firms in the portfolio and relate it to the event. This yields the average cumulative abnormal return (ACAR).

$$\mathrm{ACAR}_{\tau} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{CAR}_{it} \qquad (5b)$$

We determine the portfolio by various chracteristics of stocks such as size of the parent and quality of its product, as suggested in the hypotheses. Under the null hypothesis that each ACAR_t at time t has a mean of zero and variance of $\sigma^2_{\mathrm{ACAR}_t}$, the maximum likelihood estimate of the variance is a consistent estimator given by (Boehmer et al. 1991) as follows:

$$t_{\mathrm{ACAR}} = \frac{\frac{1}{N}\sum_{i=1}^{N}\mathrm{CAR}_i}{\sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}\left(\mathrm{CAR}_i - \frac{1}{N}\sum_{i=1}^{N}\mathrm{CAR}_i\right)^2}}. \qquad (6)$$

Data

This section briefly describes the procedure and measures of our data.

Procedure

The sample for this study consists of firms whose products were reviewed in *The Wall Street Journal* between 1991 and 2001. We collected all reviews of product quality in *The Wall Street Journal*, from the start of the feature in October 1991 until the start of this study in December 2001. Walter Mossberg authored the features, except occasionally, when he was on vacation. We collected the event dates and the name of the firm whose products were reviewed. We converted the reviews into a quality score, as explained in the next section. We then merged these data with data on returns and size of firms from the Wharton Centre for Research in Security Prices (CRSP) and Compustat data.

Measures

This subsection explains the measures for quality, events, returns, and firm size.

Quality. We define quality as a composite of attributes of which all consumers prefer more to less. Examples of such attributes are ease of use, reliability, speed, high resolution, etc. Several studies in marketing have used such a measure of quality (e.g., Archibald et al. 1983, Tellis and Wernerfelt 1987, Ratchford 1980). Although heterogeneity in consumer preferences for these attributes could create a problem in arriving at a composite, authors have shown that this situation is rare (Curry and Faulds 1986, Kopalle and Hoffman 1992). Thus, published reviews of quality seem reasonable. If heterogeneity were a problem, it would lead to support for Hypothesis H_{01} and against alternate Hypothesis H_{A1}. Anyway, we also test for the relationship between quality and its dimensions.

Because *The Wall Street Journal* does not publish numerical ratings of quality, we used a content analysis of reviews to arrive at numerical ratings, following an approach by Tellis et al. (2007). For the content analysis, we first developed a set of terms that reviewers use to describe these products. We then grouped these terms into five levels, expressing increasing quality, on a five-point scale ranging from 2 to 10 (see Appendix 1).

We recruited and trained two MBA students as research assistants (RAs) for the study. They evaluated the language of each review with the scale shown in Appendix 1. They then converted the review into a numerical value of quality. We instructed them to treat the quality scale as continuous from 2 to 10.

The first RA identified 849 reviews, and another RA independently identified 843 reviews. When pooled together, the two RAs had 765 reviews in common. This difference between the RAs comes because each had to use judgment to decide whether a product was reviewed or merely mentioned.

Events. Each RA also identified the date of the review and the name of the firm whose product was reviewed. Each recorded review constitutes an event. A third RA compiled a separate database containing the date of the review, the name of the firm, and the ticker symbol for the firm whose product was reviewed. This database contained 498 ticker symbols for the various firms whose products were reviewed. The number is less than the 765 common reviews identified by the two RAs primarily because many firms whose products were reviewed did not have ticker symbols, as they were privately held.

Returns. We used the Center for Research in Security Prices (CRSP) database available through the Wharton Research Data Services (WRDS) to collect the daily returns for the different stocks whose products were reviewed. The ticker symbols and the dates from the first data set were used to extract returns.

We extracted daily stock returns for each firm in the sample and the CRSP equal-weighted index returns. As explained earlier, for the calibration period we used a window of 180 days prior to the event. We had to remove 33 events because the price data ended before the review. The price data ended before the event because of mergers, acquisitions, delisting, and failures of the target firm. We had to remove 44 events because the price data appeared after the review; that was because the target firm's stocks were floated after Mossberg's review appeared. Thus, the final sample size was 421. The drop in sample size is due entirely to market conditions. There are 129 different firms in the sample. Many firms have a large portfolio of products with multiple new products that were reviewed by Mossberg.

The abnormal returns are from Equation (4). In all subsequent analyses, the returns are based on the closing price of that particular day. All the data are available directly from the Wharton website (http://wrds.wharton.upenn.edu). All our results are based on using this program.

To study the effects of ratings of quality on the abnormal returns, we merged the data of returns from the Wharton database with ratings of product quality we collected.

Firm Size. We used annual sales to measure the size of the firms whose products were reviewed. To do so, we extracted the annual sales data for the firm in the year their product was reviewed. The source for the sales data is the annual industrial file of the Compustat data set available through Wharton Research Data Services.

Results

This section presents results in three subsections: measurement of quality, effects of quality, and tests of hypotheses.

Measurement of Quality

Recall that our measure of quality is the coding by two RAs of Mossberg's review using the scale in Appendix 1. The correlation coefficient for the 765 reviews with common coding by the two RAs is 0.77. The interrater reliability as measured by coefficient kappa is 0.66. The RAs agree exactly on 61% of the cases and are within one count of each other for 87% of the cases. This compares favorably with the average value of 80% by Chandy et al. (2001). Thus, we use the average of the codings of the two independent RAs for subsequent analysis. In subsequent analysis we refer to this measure as overall quality, to distinguish it from the dimensions of quality we measured next.

Overall quality may be considered a function of underlying dimensions that characterize the product (Curry and Faulds 1986, Kopalle and Hoffman 1992). We sought to identify these dimensions of quality and capture the rating of the products on these dimensions. For this purpose, we hired a fourth RA whom we instructed to collect all phrases used to describe the quality of the products in every review. Based on those phrases, the prior literature, and our expertise in this area, we identified the following dimensions of quality:

Stability (crash proofness)
Compatibility (with earlier versions of the product and other products)
Ease of use
Reliability (freedom from physical breakdowns)
Utility of secondary features
Intrinsic performance (speed, clarity, resolution, etc.)
User-friendly design (simplicity, unity, parsimony).

Note that we needed a reasonably comprehensive list of dimensions (to ensure completeness) but also one that was reasonably parsimonious (to avoid too many 0s from products not being evaluated on those dimensions). To capture the evaluation of the products in these dimensions, we used an 11-point scale (0–10). In particular, we asked an RA to convert the review of the product on the above phrases into a score on the 11-point scale of the relevant dimension. For reviews that did not say anything about a product on any one of these dimensions, we substituted a value at the average of the scale (i.e., 5). For this stage of the analysis, we worked on the original 765 reviews for which we had quality information, even if a review did not have the stock market information. Some reviews were too sparse to determine individual quality dimensions, and others incorporated the value of price and popularity in the overall rating of quality. We removed these reviews from the sample. This step reduced the sample size from 765 to 733.

To investigate the correlation between the dimensions and to check for the presence of higher-order dimensions, we factor analyzed the data. We found that the dimensions were relatively orthogonal, as no dominant factors emerged.

To assess how these dimensions relate to overall quality, we regressed the overall quality score on the dimensions, using the following model:

$$\begin{aligned} Quality_i = {} & \gamma_0 + \gamma_s Stability_i + \gamma_c Compatibility_i \\ & + \gamma_e Ease\text{-}of\text{-}use_i + \gamma_r Reliability_i + \gamma_u Utility_i \\ & + \gamma_p Performance_i + \gamma_d Design_i + \mu_i, \end{aligned} \tag{7}$$

where the γs are coefficients to be estimated for each of the corresponding dimensions of quality and the

Table 2 Regression of Overall Quality on Its Dimensions

Variables	Coefficients	Std. error	t-value	Pr(>|t|)
Intercept	−4.200	1.073	−3.914	0.000
Stability	0.060	0.149	0.405	0.343
Compatibility	0.604	0.097	6.200	0.000
Ease of use	0.394	0.051	7.726	0.000
Reliability	0.140	0.091	1.541	0.061
Utility of features	0.393	0.026	14.850	0.000
Performance	0.246	0.034	7.325	0.000
User-friendly design	0.216	0.038	5.631	0.000

Note. R-square $= 0.43$; $F(7, 725) = 76.53 (p < 0.0001)$; $n = 733$.

μ_is are error terms initially assumed to identically and independently follow a normal distribution (IID normal). The results are in Table 2.

The results show that the dimensions of quality explain a substantial portion of overall quality (R^2 is 43%). All the dimensions have effects that are in the expected direction. Moreover, five of the seven dimensions of quality have effects that are significantly different from 0. The size and significance of the coefficients do not change much between the simple regression and the multiple regression. These results confirm that multicollinearity is not a problem in the data as indicated by the factor analysis. In effect, what it means is that the dimensions represent relatively independent aspects of quality and that Mossberg has been able to evaluate products independently on each dimension without suffering from any halo bias from other dimensions. All these results provide strong support for Hypothesis H_{A1}.

Effect of Quality on Returns

We first explain the models specification and then the results

Model Specification. Based on our hypotheses, we test the effect of quality on ACAR, the Average Cumulative Abnormal Returns, as defined in Equation 5b, through the following regression model:

$$\begin{aligned}\mathrm{ACAR}_{i5} = {} & \delta_0 + \delta_1 Size_{it} + \delta_2 Quality_i \\ & + \delta_3 Inferior\ Quality_i + \delta_4 Size_i * Quality_i \\ & + \delta_4 Large\ Size_i * Inferior\ Quality_i + v_i, \end{aligned} \quad (8)$$

where

$\mathrm{ACAR}_{i5} =$ average cumulative abnormal return of firm i on day 5

$Size_i =$ logarithm of sales of firm i

$Large\ Size =$ indicator variable if sales > 1bn, 0 otherwise

$Quality_i =$ rating of quality from reviews of firm i's product at time t

$Inferior\ Quality_i =$ 0 if $Quality > mean\ of\ Quality$, and 1 if $Quality < mean\ of\ Quality$

δs are coefficients to be estimated

ν_is are errors terms initially assumed to be IID normal.

In the interests of parsimony, we do not model the interaction effects of each dimension of quality on abnormal returns. Instead, we can compute these effects by multiplying the coefficients of δ in Equation 8 with the γ for the relevant dimension of quality in Equation 7. Given the possibility of correlated errors between Equations 7 and 8, we can treat them as a system of equations and estimate them via three-stage least squares (3SLS).

Results. The results of estimating Equations 7 and 8 are in Table 3.

The results show that the individual dimensions of quality have a strong and significantly positive effect on overall quality (Columns 2 and 3), and through overall quality on abnormal returns (Columns 4 and 5). This result justifies our use of the global measure of quality, in addition to the dimensions. Substantively, the result suggests that markets respond to the measures of total quality rather than to its individual components, especially as the latter are not all positively correlated with each other and may sometimes involve tradeoffs. Consistent with our expectations or informal hypotheses, the dimensions that have relatively the strongest effects on abnormal returns (in terms of coefficients or t-values) are utility of features, ease of use, and compatibility. Also, consistent with expectations, reliability seems less important than the other variables. In addition, performance and user-friendly design are important, but stability is not. The probable reason for the latter effect may be that new products have become increasingly stable in

Table 3 Regression of Abnormal Returns Versus Size and Quality

(1)	(2)	(3)	(4)	(5)
Dependent variables	Quality		ACAR	
Variable	Estimate	t	Estimate	t
Size			0.07	2.07**
Quality			0.12	2.96**
Inferior quality			0.06	1.88*
Quality × size			−0.01	−2.07**
Inferior quality × large size			−0.01	−0.25
Stability	−0.28	−0.95		
Compatibility	0.30	1.77*		
Ease of use	0.25	3.44***		
Reliability	0.10	0.86		
Utility of features	0.28	7.19***		
Performance	0.14	2.81**		
User-friendly design	0.11	1.86*		
Adj. R-square (%)	19.73%		3.52%	
F	15.75***		3.76***	
N	421		421	

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.001$; Correlation of error terms $= -0.601$.

recent years. All these results provide managers with a means of identifying what dimensions of quality to emphasize in their manufacturing and promotion. Given costs for improving quality on these dimensions, firms can use the direct effects of the dimensions of quality on ACAR as a means of evaluating returns to investing in such dimensions. Thus the framework of analysis we propose can be a means of assessing the reward or value of quality to firms.

Note that there is a decrease in R-square in the regression of quality on its dimensions from the multiple regression (Table 2) to the 3SLS estimation (Table 3). The reason is that the former regression is on a larger data set of 733 observations, which has higher variation in measures of quality and its dimensions than the latter estimation on 421 observations. We had to use the 421 observations for the latter, because for the estimation of abnormal returns (ACAR), stock market data were available for only the 421 observations.

Tests of Other Hypotheses

When discussing the tests of the other hypotheses, to provide a better understanding, in addition to the models' results above, we also display graphical results with quality and size measured categorically. Based on the above results, to obtain a purer measure of quality, we used the predicted value of quality from the regression of quality on its dimensions (Equation (7)). In so doing, we are asserting that quality that cannot be explained by its dimensions constitutes errors from sampling, observation, or measurement. (Unless otherwise stated, for the remainder of the paper, we use the term "quality" to mean the predicted quality from Equation 7 and the term "composite quality" to mean the raw quality measured from the reviews.) To graphically display the results, we collapsed quality into three categories, as follows:

Poor: Ratings of poor or unacceptable quality on our scale in Appendix 1 ($n = 33$)

Fair: Acceptable quality on our scale of quality in Appendix 1 ($n = 343$)

Good: Good and excellent quality on our scale of quality in Appendix 1 ($n = 45$).

Similarly, we collapsed firm size into two levels: small firms have sales below \$1 bn ($n = 98$) and large firms have sales above \$1 bn ($n = 323$).

We chose \$1bn as the cut-off to get a reasonable sample size for small firms. Firms reviewed were mainly very large firms, and many of the small firms that were reviewed were privately held firms with no stock prices. We also ran a multivariate regression analysis using size as a continuous variable.

Effect of the Review. This section presents the effect of the review itself, *without* taking into account the quality described in the review. This analysis ensures that the data do not contain preexisting effects or biases that are not due to quality. Recall that our dependent variable is ACAR, the Average Cumulative Abnormal Returns, as defined in Equation (5b).

Figure 3 ACAR for All New Products

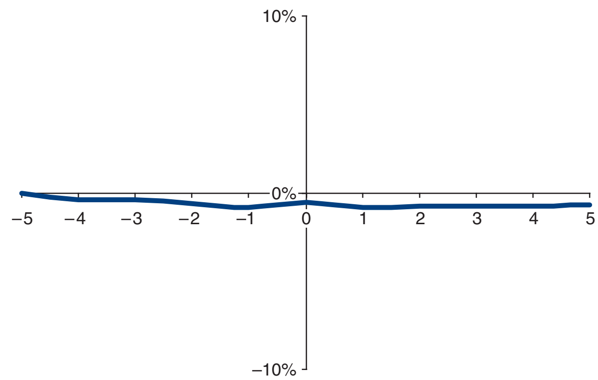


Table 3 shows that the ACAR for the firms whose new products are reviewed for a period of 5 days before and 5 days after the appearance of the review. Figure 3 shows the same effects graphically. Note that all effects are below 0, but most are not significantly different from 0. Recall that this analysis is for *all* products, *without* taking quality into account. The result suggests that our sample does not contain stocks with any unusual properties. It also shows that the mere publication of a review by Walter Mossberg, without taking quality into account, does not affect stock returns abnormally. These results are tabulated in Table 4.

We next proceed to ascertain how the quality of the review affects the abnormal returns of the firms whose products are reviewed.

Effect of Quality. Hypotheses H_{A2} states that abnormal returns will vary with the rating of quality. In other words, the more positive the rating of a product's quality, the higher the firm's abnormal returns will be. Columns 4 and 5 of Table 3 show that the main effect of quality on ACAR is positive and significantly different from 0. Similarly, Columns 2 and 3 of the same table show that the effects on ACAR

Table 4 ACAR for All Products

Day	ACAR(%)	t-value
−5	0.0	0.00
−4	−0.4	−1.32
−3	−0.4	−1.23
−2	−0.6	−1.66
−1	−0.8	−1.94
0	−0.8	−1.94
1	−0.6	−1.26
2	−0.9	−1.71
3	−0.8	−1.42
4	−0.8	−1.44
5	−0.8	−1.32

Note. $n = 421$.

Figure 4 ACAR by Quality

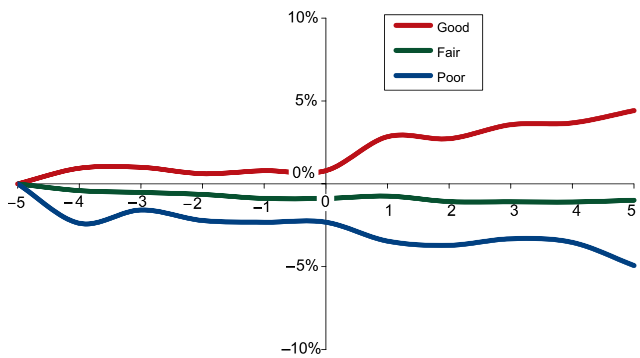
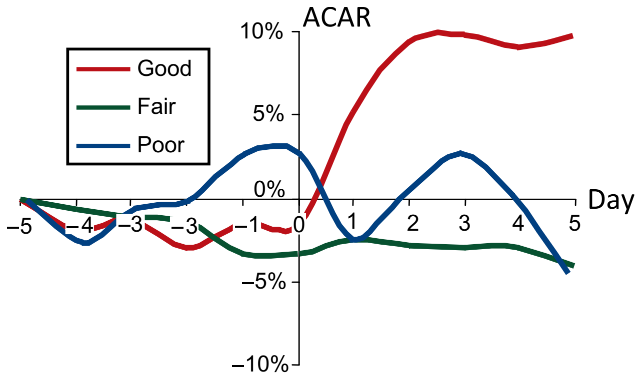


Figure 5 ACAR for Small Firms by Quality



of five of the seven dimensions of quality are positive and significantly different from 0. These results provide strong support for Hypothesis H_{A2}.

The graphical analysis also bears out these findings. Figure 4 shows the average cumulative abnormal returns for three levels of quality for a window of 5 days before the event and 5 days after the event.

ACAR of firms with poor quality decline steadily throughout the event window. ACAR for firms with good quality rises on the day of the review and continues to rise consistently for about 5 days. In the interest of parsimony we keep the postevent window fixed at 5 days for the rest of the analysis. Table 5 shows the statistical tests for these results. It shows that the ACAR of fair and good products at 5 days after the event are superior to those of poor-quality products. The differences are significantly different from 0 for the comparison of good relative to poor. These results also support Hypothesis H_{A2}.

Negativity Bias. H_{A3} states that information about unfavorable quality affects returns more than information about favorable reviews. Note that we define a variable, *Inferior Quality*, which takes on value 0 for *Quality* above its mean and 1 for quality below its mean. Columns (4) and (5) of Table 3 show that the coefficient of *Inferior Quality* is positive, over and above the effect of quality, which is included in the model. This result does not support H_{A3}. However, the graphical analyses directionally support H_{A3}. Figure 4 shows that by day 5, the ACAR of firms with poor quality falls a little more than the ACAR of firms with good quality rises. The ACAR for firms with poor reviews at day 5 is −4.93%, while that for firms with good reviews 4.4%.

Table 5 Difference in ACAR by Quality at Five Days

Quality difference	Difference in ACAR (day 5) (%)	*t*-value
Good – Poor	9.35	19.52*
Fair – Poor	3.94	4.59*

Notes. $n = 45$ (good); 343 (fair); 33 (poor).
*$p < 0.001$.

Reputational Asymmetry. Hypothesis H_{A4} states that because of reputational asymmetry, inferior quality leads to more negative abnormal returns for large firms than for small firms, and superior quality leads to more positive abnormal returns for small firms than for large firms.[1] Table 3 shows that the coefficient for *Quality ∗ Size* is negative and significantly different from 0. Consistent with H_{A4}, this result implies that higher quality (over and above the main effect of quality) translates into lower returns for larger firms than for smaller firms. Further support for H_{A4} comes from the coefficient of *Inferior Quality ∗ Large Size* (both dummy variables), which is negative. Note that both formats of the interaction terms (continuous and dummy) were in the expected direction. Thus, we retained both.

We can see this effect by plotting abnormal returns across quality separately for large and small firms.

Figure 5 shows that ACAR for small firms whose products are rated good shoots up to 5.11% on the day of the announcement and to around 10% by day 5. In contrast, the ACAR for small firms with poor quality fluctuates around 0 and declines to about −4.6% by day 5. The market seems to be more surprised by and thus better rewards the good quality of small firms than it penalizes poor quality of small firms.

Figure 6 shows that the ACAR for large firms with good quality rise to about 2.3%, and those for fair quality are approximately 0 by day 5. However, the ACAR for large firms with poor quality keep falling and stabilize at a large negative value—around −5%—by day 5. Thus, the market seems to be surprised by and more heavily penalizes the poor quality

[1] Although Table 3 shows a small significant effect for size, that effect is an artifact of the multiple interaction terms with size. A separate regression without the size interactions shows that the main effect of size is not different from 0.

Figure 6 ACAR for Large Firms by Quality

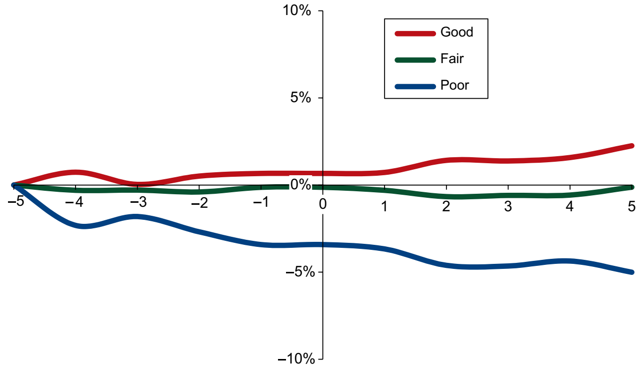
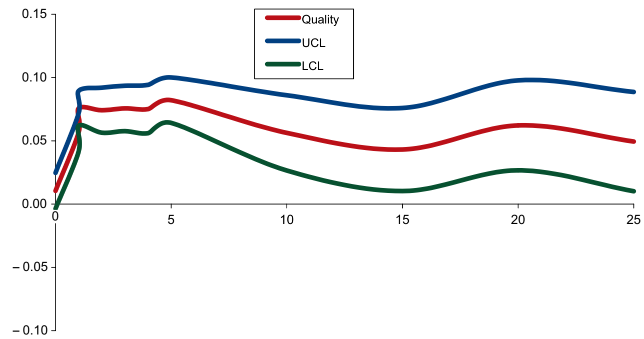


Figure 7 Impact of Quality Rating on Mean Returns (with 95% Confidence Intervals)



of large firms than it rewards the good quality of large firms.

To test the reputational asymmetry in this context, we compare the difference in abnormal returns of large firms with small firms, both with good quality. Table 6(a) tests the difference in these ACAR values. We subtracted the ACAR of large firms from those of small firms, both with good quality.

Table 6(a) shows that the difference in ACAR for firms with good quality is higher for small firms than for large firms for all the days in the event period. Also, the differences are highly significant on all days. In an analogous manner, Table 6(b) shows the differences in ACAR of large firms from those of small firms, both with poor quality.

Long-Run Analysis. To study the long-run implications of the returns to quality we looked at how the coefficient for quality behaves over time. For this analysis we extended the regression on abnormal returns to 25 days. To visually demonstrate the multivariate analysis, we run Equation (7) for each day from day −5 to day 25, for quality and size measured continuously, as above. Figure 7 plots the regression coefficients along with their upper and lower confidence intervals, from day −5 to day 25. The figure shows the dramatic increase in returns for reviews of higher quality beginning with day 0. The results remain positive over time, but the confidence widens because of the increase in noise from extraneous events as we move further from the focal event.

The numeric results are in Table 7.

Table 6a Difference of ACAR for Good Quality

Days	Mean difference (%)	*t*-value
1	2.81	2.83*
5	7.57	7.74*

Notes. n (large firms) = 31; n (small firms) = 14.
*$p < 0.001$.

Table 6b Difference of ACAR for Poor Quality

Days	Mean difference (%)	*t*-value
1	−1.15	−1.465
5	−0.41	−0.42

Notes. n (large firms) = 27; n (small firms) = 6.

Discussion

Quality is an important marketing and strategic variable. Yet, it tends to be undervalued by firms, probably because it is difficult to measure. Computing the abnormal returns to quality reviews of new products provides a measure of how financial markets value the quality of those new products. To the extent that such markets are efficient, the results provide us with a new and powerful metric for the assessment of quality. The abnormal returns that we detect can also be considered as a quality premium paid by investors.

The major problem in carrying out such an analysis is to find a consistent and systematic source of information on the quality of new products that would be accessible to investors. We propose that reviews of the quality of new products published in *The Wall Street Journal* are one such source. We analyzed the abnormal returns of quality measured in these reviews from 1991 to 2001.

Theory and the literature suggest several null and alternative hypotheses about how the abnormal

Table 7 Quality Coefficients Day 0 to Day 25

Day	Quality	*t*-value
0	0.057	3.572
1	0.076	4.277
2	0.074	3.968
3	0.076	3.990
4	0.075	3.736
5	0.082	3.819
10	0.056	1.799
15	0.043	1.250
20	0.062	1.663
25	0.049	1.196

Note. $n = 421$.

returns of new products would vary by quality and firm size. Our results show strong support for most hypotheses. Most important, we find that a poor review on quality hurts firms by leading to substantially and significantly lower abnormal returns than for some good reviews. In addition, there is some negativity bias, in that negative returns to poor reviews of quality are greater in absolute value than positive returns to good reviews. There is also a small reputational asymmetry in that rewards to small firms with good reviews of quality are greater than those to large firms with good reviews. On the other hand, large firms are penalized more by poor reviews of quality than they are rewarded for good reviews.

The results raise five important questions:

How strong are these effects?

How do returns for new products compare with those for reviews of quality of new products?

How stable are these effects over time?

How robust are these results?

Is there leakage of information before the event?

Strength of Effect

The effects that we identify are very strong relative to those obtained for comparable studies in marketing (see Table 1). Those studies show that the effect in past studies varies from almost 0% to 0.72%. In contrast, we find that the five-day ACAR ranges from 9.83% for small firms with products with good reviews of quality to −5% for large firms with products of poor review of quality. Our results also compare well with a study done by *Wired* magazine. It showed that stocks of firms with positive reviews went up an average of 14%, but one with poor reviews went down 9% (see Table 8 for details). Whereas The *Wired* results are from a selective sample, but our results are based on a systematic and formal analysis.

Table 8 **Returns for Four Product Reviews, by *Wired* Magazine**

Walt said	What happened
CNET is the real deal—a serious, scrappy online news organization (1996)	CNET's stock climbed 33% in the five days after the review.
This is simply the most gorgeous personal computer I've ever seen or used (2000)	Apple's stock price jumped 10% the day after his Cube review hit stands.
The initial radios are poorly designed ... Navigation is a nightmare (2002)	XM's stock dropped nearly 9% the day Mossberg panned its radios.
Overall, we like Vialta's idea of incorporating your television into a videophone (2003)	Vialta's stock rose 15% the day the column ran.

Source. Adapted from Deutschman (2004).

New Products vs. Quality

Firms are under pressure to rush to bring new products to market. In so doing, they might overlook the risks of marketing new products and sacrifice quality (Tellis and Golder 2001). We could not directly assess the trade-off between a strategy of rushing out new products with inferior quality and waiting to introduce new products of superior quality. However, we can get an indirect assessment of this trade-off by contrasting the returns for quality reviews in our study to those for new products in the literature. Chaney et al. (1991) found that cumulative abnormal returns for *all* new products were on average 0.82% one day after the announcement of the new product in *The Wall Street Journal*. In contrast, we find that firms with poor quality reviews suffer a drop in ACAR of about 5% five days after the review appears in *The Wall Street Journal*. In comparison, firms with good quality reviews enjoy a gain of about 10% over the same period. Thus, the premium for good quality and the penalty for poor quality are several times higher than that for an average new product as measured by Chaney et al. (1991). Hence, managers are well advised to wait to ensure that their products are of good or very good quality before introducing them.

Overall, products earn more positive than negative reviews, so firms are not all behaving suboptimally by bringing inferior products to the market. At the same time, only a minority of firms earns very good reviews of quality. Thus, most firms have room for improvement.

Stability of Results

Does the market fully realize all rewards to quality immediately on publication of the review? Alternatively, are the returns that we observe merely an overreaction that dissipates once we look over a longer window? An assessment of medium-term returns will reveal whether the five-day returns are an overreaction, underreaction, or fair reaction to quality reviews.

The literature provides no certain guide as to what constitutes the medium or long term. We judged that at most, we should consider a period of 10 days after the event. A period longer than 10 days could be considered excessive, as a large number of extraneous factors could blur the effects and confuse the interpretation of results during that period.

To check the medium-term effect of reviews on quality, we constructed buy-and-hold portfolios of firms by quality reviews, following an approach first used by DeBondt and Thaler (1985). We formed three portfolios of products based on their reviews (good, fair, and poor). We added a firm to the portfolio on the date its review appeared. We tracked the market-adjusted daily returns of these firms for 10 days after inclusion. We computed the average daily returns of each portfolio, synchronizing returns by date of publication of the review. Over an interval of two trading days t_1 and t_2 the buy-and-hold returns (BHAR) for a firm i is calculated as follows:

$$\mathrm{BHAR}_i = \left[\prod_{t=t_1}^{t_2}(1+r_t^i)-1\right] - [(1+\hat{\alpha}_i)^{(t_2-t_1+1)}-1] - \beta_i\left[\prod_{t=t_1}^{t_2}(1+r_{mt})-1\right], \qquad (9)$$

where, r_t^i, r_t^m are returns to the firm i and the market at time t, $\hat{\alpha}_i$ is idiosyncratic risk of the firm, and β_i is the correlation of the firm's returns to the market.

In effect, the compounded return of the market is taken away from the compounded return of the firm and adjusted for its individual risk. The average BHAR for any portfolio of N firms then is given by

$$\text{Average BHAR} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{BHAR}_{i,t_1,t_2}. \qquad (10)$$

The average buy-and-hold is compounded return as opposed to the cumulative return. In practice, we obtain these returns using the Wharton Eventus program available through Wharton. Figure 8 shows the buy-hold returns across quality.

Neither the effects of good quality nor those of bad quality dissipate quickly or return to 0. This result suggests that the return for five days presented earlier is not an overreaction or flash in the pan. On the contrary, the returns for good reviews and poor reviews extended the five-day trend in returns for almost 10 days. Returns beyond 10 days show a small, steady increase in the same direction; however, these results are not significantly different from 0. The reason, consistent with the efficient market hypothesis, is that as we go past the event, multiple extraneous events increase the variance in returns and reduce the possibility of finding any effects that are significantly different from 0.

Figure 8 **Buy-Hold Returns Across Levels of Quality**

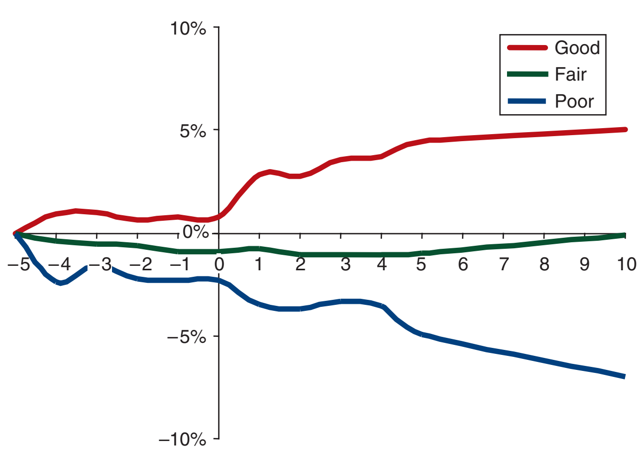


Table 9 **Sensitivity of Returns of Firms with Good Quality Ratings to Market Index**

Day	CRSP (%)	NYSE (%)	NASDAQ (%)
−5	0	0	0
1	2.84	3.20	2.84
5	4.42	6.01	2.64
10	5.05	5.84	3.74

Robustness Checks

To check the robustness of our results we checked the sensitivity of returns to alternate market indices, the power of the overall study, and the variation of returns to exogenous factors.

First, we looked at how the returns compared across three alternate indices for market returns when computing abnormal returns: NASDAQ, NYSE, and CRSP. Table 9 shows the ACARs for good quality firms using each of these indices. The results show that although the CRSP and NYSE track one another, the NASDAQ is slightly lower than the two.

Second, we investigated the power of our event study to detect abnormal returns. This is the power to reject the null hypothesis that the abnormal return is 0. Given our sample sizes of 98 small firms and 323 large firms and the size of the abnormal returns (9.8% to −5.0%), the power to detect the abnormal return is 1 for a test with a size of 5% (Campbell et al. 1997, p. 170).

Third, we also analyzed variation in returns by (a) type of product, (b) date of product introduction, and (c) timing of event relative to the dot.com bust (in March 2000). However, we found no significant effects for any of these analyses.

Fourth, we conducted an outlier analysis by deleting 2% and 5% of the extreme returns in the sample, but these deletions did not affect the results.

All these tests suggest further confidence in our main findings.

Information Leakage

The decline in abnormal returns for large firms with poor quality start before the review is published, raising the possibility of leakage of information. However, we do not think that this effect is actually caused by leakage of information, because if it were, it would occur for negative and positive reviews and for small and large firms. Rather, the decline in abnormal returns before the event for large firms with inferior quality may occur for several other reasons. First, reviews in other outlets may cause negative publicity, especially because large firms are more widely covered and expected to have good products; their poor performance is likely to lead to negative surprise. Second, large firms with poor quality may be suffering an erosion of market share from mean reversion that could negatively affect returns. Third, such firms may have increasing costs from increasing competition, which negatively affects their returns. Our design does not enable us to test these rival explanations. However, this limitation does not compromise all our other results, because those results are consistent with theory.

Implications

This study has six important implications for research, strategy, investing, and policy.

First, researchers have an alternate metric by which to assess the markets' evaluation of the quality of products—the ratings of quality in published reviews. Though event studies do not imply causality, abnormal returns are one measure of valuation when we do not have other measures, such as return on investment. This is because accounting data are backward looking and rely on past data, but financial markets are forward looking and incorporate expectations. In effect, the abnormal returns to unanticipated information such as quality reviews are the verdict of millions of investors in the future worth of the firm as the result of this new information. Moreover, our study shows that the dimensions of quality that directly affect product attributes when translated into a suitable quantitative scale can lead to testable effects in terms of abnormal returns to the stocks of the parent firms.

Second, from a strategic perspective, merely introducing new products is no guarantee of future success or improved stock market performance. It is essential that firms introduce new products that are of some minimum quality. Indeed, it might be difficult to recover from a negative review of quality. Thus, a short-term management focus may lead to a serious undervaluation of quality.

Third, the results of regression of *Overall Quality* on its dimensions may indicate factors that strategists and new product designers need to consider. Based on partial R^2 and t-stats, we found that compatibility, performance, ease of use, and utility of features were particularly important. The probable reason was that new products are coming out with increasing frequency and consumers want to ensure that these products perform well, are easy to use, and are compatible with their old products. On the other hand, reliability and stability were relatively less important. The probable reason is that new high-tech products are often replaced just before the expiration date of their warranty, so the latter two attributes are no longer that critical.

Further, the archive of past ratings may indicate common errors that marketers of new products commit. Our qualitative assessment of this archive indicates some of the common problems in the quality of new products. In particular, characteristics that get a poor review are inconvenient and confusing to use or have a complicated design (see Table 10).

Table 10 Major Reasons for Poor Rating

Reasons	Frequency
Inconvenient and confusing to use	11
Poor/complicated design	12
Instability/poor reliability	5

Fourth, careful investors can use these results to profit by disinvesting in firms whose quality receives a very poor review in *The Wall Street Journal*. However, it is quite likely that once this pattern of abnormal returns has been validated and becomes public knowledge, investors' actions may cause it to disappear, as was the case with many other findings (Horowitz et al. 2000).

Fifth, our study shows that published reviews of new products provide a valuable service to the investing public by informing them about the prospects of new products and the firms that introduce them.

Sixth, managers of *The Wall Street Journal* must ensure that their team of reviewers continues to not exploit the trading opportunity for personal profit, as seems to be the case so far.

Limitations and Future Research

This research has several limitations, some of which could serve as areas for future research. First, event study analyses usually reflect shareholder perceptions. Although these might be, on average, correct, they need not always reflect consumer perceptions. Second, event analysis does not imply causality between quality ratings and abnormal returns. Mossberg's review may just be increasing the salience of the firm for the investing public. Third, we did not explicitly record the announcement of the new products by parent firms, as discussed above. Fourth, our study considers only electronic products and software. It would be interesting to test the generalizability of the results to other product categories.

Acknowledgments

The authors thank Dan Ariely for suggestions about the analysis and Ashish Sood for help on some data analysis. This paper benefited from comments at seminars at the University of Florida, Ohio State University, and the University of Maryland.

Appendix 1

(Adapted from Tellis et al. 2007)

Content Analysis Outline

The outline for quantifying review information is given as follows:

(1) Excellent—10: A market leader that offers exceptional performance

- It is considered the most powerful product available today.
- This product is the big winner.

- Editor's Choice.
- This product is excellent.
- This product could be one of those milestones that change the way we use computers.
- It is unquestionably the most powerful product you can buy.
- It is miles ahead of the competition.
- The product stands at the top.
- It is the very best product of the year.
- This product has a very good chance of establishing a new standard.
- It is one of the products that does everything right.
- It is clearly the most richly endowed product that you can purchase.
- It is an outstanding performer for its wealth of features and flexibility.

(2) Good—8: Excels in many areas; a good buy

- This product is an attractive alternative.
- This product is a good choice.
- This product is a serious threat to the current standard.
- It is an impressive product.
- It is a richer product than its principal competitors.

(3) Acceptable—6: Average for its class; a justifiable purchase

- The product is well thought out, but there are still a few problems with it.
- It is an economical and elegant program. Is it the right product for you? As usual, it depends.
- It is a popular choice. However, it may not make you happy.
- It is a strong competitor to its rival. However, its major weakness is . . . .

(4) Poor—4: Out of date or substandard; positives offset by more negative features

- It is a product I would love to love, but can't.
- It has been outdistanced by its competitors.
- It looks dim beside its competition.
- In many ways, it still clings awkwardly to its past.
- It performs unsatisfactorily.

(5) Unacceptable—2: Missing necessary features; avoid

- It scored the lowest in overall satisfaction.
- It occupies the lowest spot.
- It is definitely bad.
- It is very poor.
- It performs quite sluggishly.
- Definitely avoid/do not buy.

Appendix 2. Beta Values at the Portfolio Level

	Estimate based on 25 days	
Quality level	Before event	After event
Poor	1.9*	1.98*
Fair	1.9*	1.96*
Good	1.78*	1.81*

*$p < 0.001$.

References

Aaker, D. A. 1995. *Building Strong Brands*. The Free Press, New York.

Aaker, D. A., R. Jacobson. 1994. The financial information content of perceived quality. *J. Marketing Res.* **31**(May) 191–201.

Agarwal, J., W. A. Kamakura. 1995. The economic worth of celebrity endorsers: An event study analysis. *J. Marketing* **59**(July) 56–62.

Archibald, R. B., C. A. Haulman, C. E. Moody, Jr. 1983. Quality, price, advertising, and published quality ratings. *J. Consumer Res.* **9**(March) 347–356.

Ball, R., P. Brown. 1968. An empirical evaluation of accounting income numbers. *J. Accounting Res.* **6**(2) 159–178.

Boehmer, E., A. Poulsen, J. Musumeci. 1991. Event study methodology under conditions of event-induced variance. *J. Financial Econom.* **30** 253–272.

Buzzell, R., F. Wiersema. 1981. Modeling changes in market share: A cross-sectional analysis. *Strategic Management J.* **2**(January–February) 27–42.

Buzzell, R., B. Gale, R. Sultan. 1975. Market share—A key to profitability. *Harvard Bus. Rev.* **53**(January–February) 97–106.

Campbell, J. Y., A. W. Lo, A. C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Carpenter, G. S., K. Nakamoto. 1989. Consumer preference formation and pioneering advantage. *J. Marketing Res.* **26**(3) 285–298.

Chandy, R. K., G. J. Tellis, D. J. Macinnis, P. Thaivanich. 2001. What to say when: Advertising appeals in evolving markets. *J. Marketing Res.* **38**(November) 399–414.

Chaney, P. K., T. M. Devinney, R. S. Winer. 1991. The impact of new product introductions on the market value of firms. *J. Bus.* **64**(4) 573–610.

Curry, D., D. Faulds. 1986. Indexing product quality: Issues, theory, and results. *J. Consumer Res.* **13** 134–143.

DeBondt, W., R. Thaler. 1985. Does the stock market overreact? *J. Finance* **40** 793–805.

Deutschman, A. 2004. The kingmaker. *Wired* (May), http://www.wired.com/wired/archive/12.05/mossberg.html?pg=1&topic=mossberg&topic_set.

Dodds, W. B., B. K. Monroe, D. Grewal. 1991. Effects of price, brand, and store information on buyers product evaluations. *J. Marketing Res.* **28**(August) 307–319.

Dolley, J. 1933. Characteristics and procedure of common stock split-ups. *Harvard Bus. Rev.* **11** 316–326.

Einhorn, H. J., R. Hogarth. 1978. Confidence in judgement: Persistence in the illusion of validity. *Psych. Rev.* **85**(5) 395–416.

Eliashberg, J., S. M. Shugan. 1997. Film critics: Influencers or predictors? *J. Marketing* **61** 68–78.

Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *J. Finance* **25**(2) 383–417.

Fama, E. F. 1991. Efficient capital markets: II. *J. Finance* **46** 1575–1617.

Fama, E. F. 1998. Market efficiency, long term returns and behavioral finance. *J. Financial Econom.* **49** 283–306.

Fama, E., L. Fisher, M. Jensen, R. Roll. 1969. The adjustment of stock prices to new information. *Internat. Econom. Rev.* **10** 1–21.

Geyskens, I., K. Gielens, M. G. Dekimpe. 2002. The market valuation of Internet channel additions. *J. Marketing* **66** 102–119.

Golder, P. N., G. J. Tellis. 1993. Pioneering advantage: Marketing logic or marketing legend. *J. Marketing Res.* **30**(2) 158–170.

Hjorth-Andersen, C. 1984. The concept of quality and the efficiency of markets for consumer products. *J. Consumer Res.* **11**(September) 708–718.

Horowitz, J. L., T. Loughran, N. E. Savin. 2000. The disappearing size effect. *Res. Econom.* **54** 83–100.

Horsky, D., P. Swyngedouw. 1987. Does it pay to change your company's name? A stock market perspective. *Marketing Sci.* **6**(4) 320–335.

Houston, M. B., S. A. Johnson. 2000. Buyer-supplier contracts versus joint ventures: Determinants and consequences of transaction structure. *J. Marketing Res.* **37** 1–15.

Ittner, C., D. F. Larcker. 1997. Product development cycle time and organizational performance. *J. Marketing Res.* **34**(February) 13–23.

Jacobson, R., D. Aaker. 1987. The strategic role of product quality. *J. Marketing* **51**(October) 31–44.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**(March) 263–291.

Kalyanaram, G., W. T. Robinson, G. L. Urban. 1995. Order of market entry: Established empirical generalizations, emerging empirical generalizations and future research. *Marketing Sci.* **14**(3) 212–221.

Kamakura, W. A., B. T. Ratchford, J. Agarwal. 1988. Measuring market efficiency and welfare loss. *J. Consumer Res.* **25**(3) 289–302.

Koku, P. S., H. S. Jagpal, P. V. Viswanath. 1997. The effect of new product announcements and preannouncements on stock price. *J. Market-Focused Management* **2**(2) 183–199.

Kopalle, P., D. Hoffman. 1992. Generalizing the sensitivity conditions in an overall index of product quality. *J. Consumer Res.* **18** 530–535.

Kordupleski, R. E., R. T. Rust, A. J. Zahorik. 1993. Why improving quality doesn't improve quality (or whatever happened to marketing?). *California Management Rev.* **35**(Spring) 82–95.

Lane, V., R. Jacobson. 1995. Stock market reactions to brand extension announcements: The effect of brand attitude and familiarity. *J. Marketing* **59**(January) 63–77.

Liebowitz, S. J., S. E. Margolis. 1990. Fable of the keys. *J. Law and Econom.* **33**(1) 1–25.

Malkiel, B. 1992. Efficient market hypothesis. P. Newman, M. Milgate, J. Eatwell, eds. *New Palgrave Dictionary of Money and Finance*. Macmillan, London, UK, 739–742.

Mayzlin, D. 2006. Promotional chat on the internet. *Marketing Sci.* **25**(2) 155–163.

Metrick, A., R. Zeckhauser. 1998. Price versus quantity: Market clearing mechanisms are uncertain about quality. *J. Risk Uncertainty* **17** 215–242.

Mittal, V., E. Anderson, A. Sayrak, P. Tadikamalla. 2005. Dual emphasis and the long-term financial impact of customer satisfaction. *Marketing Sci.* **24**(4) 544–555.

Mitra, D., P. N. Golder. 2006. How does objective quality affect perceived quality? Short-term effects, long-term effects, and asymmetries. *Marketing Sci.* **25**(3) 230–247.

Moorman, C., R. Du, C. F. Mela. 2005. The effect of standardized information on firm survival and marketing strategies. *Marketing Sci.* **24**(2) 263–274.

Nelson, P. 1970. Information and consumer behavior. *J. Political Econom.* **78** 311–329.

Pauwels, K., J. Silva-Risso, S. Srinivasan, D. M. Hanssens. 2004. New products, sales promotions and firm value, with application to the automobile industry. *J. Marketing* **68** 142–156.

Phillips, L., D. Chang, R. Buzzell. 1983. Product quality, cost position and business performance: A test of some key hypotheses. *J. Marketing* **47** 26–43.

Ratchford, B. T. 1980. The value of information for selected appliances. *J. Marketing Res.* **17**(1) 14–25.

Robinson, W. T., C. Fornell. 1985. The sources of market pioneer advantages in the consumer goods industries. *J. Marketing Res.* **22** 297–304.

Roush, W. 2004. Technology's Uncle Walter. *Technology Review* (May 7), http://www.technologyreview.com/editors/15544.

Rust, R. T., J. J. Inman, J. Jia, A. Zahorik. 1999. What you don't know about customer-perceived quality: The role of customer expectation distributions. *Marketing Sci.* **18**(1) 77–92.

Shankar, V. 1999. New product introduction and incumbent response strategies: Their interrelationship and the role of multimarket contact. *J. Marketing Res.* **36**(3) 327–344.

Shaked, A., J. Sutton. 1983. Natural oligopolies. *Econometrica* **51**(5) 1469–1483.

Shefrin, H. 2000. *Beyond Greed and Fear*. Harvard Business School Press, Boston, MA.

Srivastava, R. K., T. A. Shervani, L. Fahey. 1998. Market-based assets and shareholder value: A framework for analysis. *J. Marketing* **62** 2–18.

Subramani, M., E. Walden. 2001. The impact of e-commerce announcements on the market value of firms. *Inform. Systems Res.* **12**(2) 135–154.

Sutton, J. 1986. Vertical product differentiation: Some basic themes. *Amer. Econom. Rev.* **76**(2) 393–398.

Tellis, G. J., C. Fornell. 1988. Advertising and quality over the product life cycle: A contingency theory. *J. Marketing Res.* **15**(1) 64–71.

Tellis, G. J., P. N. Golder. 2001. *Will and Vision: How Latecomers Grow to Dominate Markets*. McGraw-Hill, New York.

Tellis, G. J., B. Wernerfelt. 1987. Competitive price and quality under asymmetric information. *Marketing Sci.* **6**(3) 240–253.

Tellis, G. J., Y. Yin, R. Niraj. 2007. Network effects or quality: What drives the success of new high-tech products? Working paper, Marshall School of Business, University of Southern California, Los Angeles, CA.

Zhang, J., C. Narasimhan. 2000. Market entry strategy under firm heterogeneity and asymmetric payoffs. *Marketing Sci.* **19**(4) 313–327.