Review of Accounting Studies, 6, 305–329, 2001

Back to Basics: Forecasting the Revenues of Internet Firms

BRETT TRUEMAN trueman@haas.berkeley.edu
Donald and Ruth Seiler Professor of Public Accounting, Haas School of Business, University of California, Berkeley, Berkeley, CA 94720-1900

M. H. FRANCO WONG
Assistant Professor of Accounting, Haas School of Business, University of California, Berkeley, Berkeley, CA 94720-1900

XIAO-JUN ZHANG
Assistant Professor of Accounting, Haas School of Business, University of California, Berkeley, Berkeley, CA 94720-1900

Abstract. This paper examines the roles played by past revenues, web usage data, and analysts in forecasting the future revenues of internet firms during the years 1998 to 2000. For this time period our analysis shows that estimates of web traffic growth have significant incremental value in the prediction of revenues above time-series forecasts. Furthermore, analysts almost always underestimate the revenues of internet firms. Historical revenue growth has incremental predictive power over analysts' forecasts for portal and content/community firms, but not for our e-tailer sample. Moreover, the stocks of the portal and content/community firms with high historical revenue growth earn higher abnormal returns during our sample period than do those with low historical growth. Estimates of web usage growth generally do not have incremental value over analysts' forecasts for predicting the revenues of either set of firms. However, perfect foreknowledge of actual web usage growth would provide incremental predictive power. Collectively, our findings point to the potential value for forecasting purposes of both improving upon the web usage estimates and obtaining more timely reports of actual web traffic.

Keywords: analyst forecasts, forecasting, internet, revenues, web traffic

Forecasting revenues is an essential first step in the valuation of any publicly traded company. This is especially true in the internet industry, where so few firms are reporting profits and where investors and investment professionals have turned to the price/revenue ratio, rather than the price/earnings ratio, to measure relative valuations in the marketplace.[1] It is in this context that this paper examines the roles played by historical revenues, web usage data, and analysts in the prediction of the revenues of internet firms.[2]

It is by no means an easy task to forecast internet firm revenues. The internet industry, and the firms within it, are so young that there is little historical financial information available with which to generate forecasts—most of the firms in our sample have been public for two years or less. Moreover, these firms are evolving at such a rapid and unpredictable pace that past revenue numbers may have limited usefulness for forecasting purposes.

In this study we augment the historical financial numbers with non-financial data on web usage. This data come directly from the internet companies, themselves, as well as from independent rating firms (such as Media Metrix, PC Data, and Nielsen//Netratings).

It includes, among other numbers, statistics on web site pageviews, visitors, and minutes spent per page. There are several reasons to expect usage levels at an internet firm's web site(s) to be positively related to its revenues. First, higher web usage likely reflects greater online demand for the company's products and services. Second, increased traffic leads to greater revenue bookings from existing advertisers. Finally, higher usage likely attracts more advertisers and, at least indirectly, allows the firm to raise the rates charged for future ads.

Our study focuses on a subset of the internet stock universe—the portals (those providing a gateway to the internet), the content/community providers (those catering to certain segments of the population or to groups of people with specific interests), and the e-tailers (who sell goods and services on the internet). These firms share a common characteristic in that their primary business involves direct contact with users on the web. Other types of internet firms, such as those providing security or those solely offering internet access, were excluded from our study, as they are of a distinctly different nature. Our analysis spans the time period from the fourth quarter of 1998 through the second quarter of 2000.

We begin by showing that current quarterly revenue growth is significantly associated with one-quarter lagged revenue growth for the portals and content/community providers, but not for the e-tailers. The weak association for the e-tailers is likely due, at least in part, to seasonalities which are widespread in the retail industry. Current revenue growth is also shown to be significantly related to contemporaneous growth in web traffic, more so for the e-tailers than for the portals and content/community providers.

We next compare several time-series forecasting models, each based on historical revenues. This analysis is especially important when valuing firms with little or no analyst coverage, as there is, of necessity, an increased reliance on historical revenues for forecasting purposes. We find a forecast that assumes a constant change in quarterly revenue performs better overall (along dimensions such as average and mean absolute percentage error) than the other time-series models considered. Taking the constant revenue change time-series forecast as a benchmark, we then examine whether estimated growth in web traffic has incremental predictive power for quarterly revenues, by including both the time-series forecast and the estimate of quarterly web traffic growth in a regression with current quarterly revenue growth as the dependent variable. We find estimated traffic growth to have significant incremental predictive value for the e-tailers over and above historical revenues. In contrast, there is no significant incremental predictive value for the portals and content/community providers.

The focus of our analysis then turns to a subset of our sample firms—those followed by analysts–in order to examine the predictive ability of analysts' revenue forecasts. As expected, analysts' forecasts are highly correlated with realized quarterly revenues. However, in contrast to prior research for other industries, which finds an *optimistic* mean bias to analysts' earnings forecasts, analysts almost uniformly *underestimate* revenues for our internet firms. The average underestimation is over 9 percent, with *86 percent* of the consensus analyst forecasts falling below realized revenues.

We next test whether historical revenue growth and estimated growth in web usage have incremental predictive power for quarterly revenues above analysts' forecasts. If analysts fully incorporate all publicly available information in their forecasts, then these measures should not have any additional predictive ability. However, the consistent underestimation of revenues by analysts suggests that, at least for internet firms, we may find evidence of incremental predictive value. Our analysis yields mixed results. For our portal and content/community subsample we find that past revenue growth provides significant incremental predictive power for current revenue growth over and above analysts' revenue forecasts. However, this result does not extend to the e-tailer subsample. Furthermore, estimated web usage growth generally does not have significant predictive ability above analysts' forecasts for either subsample. We do find, though, that perfect foreknowledge of actual quarterly web usage growth would provide incremental predictive power for current quarterly revenues over analysts' forecasts. This points to the potential value for forecasting purposes of both improving upon the web usage estimates and obtaining more timely reports of actual web traffic.

That historical revenue growth has incremental predictive power above analysts' forecasts for the portals and content/community providers suggests that the firms with stronger past revenue growth might earn greater abnormal returns during our sample period than those with weaker past growth. Similarly, the finding that estimated web usage growth has incremental predictive ability for e-tailers' revenues above historical revenues suggests that the firms with higher estimated web traffic growth might outperform those whose growth is lower. Our analysis confirms that, during our sample period, the portals and content/community firms with higher past revenue growth earn significantly higher abnormal returns than those with lower past growth. However, no significant difference is found for the e-tailers.

The recency of the internet phenomenon restricts our data and analysis to a relatively small number of quarters. Moreover, the time period we study, characterized by both high and very variable revenue growth rates, as well as highly volatile stock returns, reflects a significant degree of uncertainty regarding internet firms' future prospects. Consequently, the relations we have documented among analysts' forecasts, historical revenues, and web traffic metrics, must be interpreted with some caution. Furthermore, because our return analysis is conducted within-sample, conclusions cannot be drawn as to the profitability of our strategies out-of-sample. Data from future quarters will provide the opportunity to test the stability of our findings over time.[3]

The plan of this paper is as follows. In Section 1 we describe our sample and the data used in the study. The relation between internet firm revenue growth and measures of web traffic growth is analyzed in Section 2. In Section 3 we compare several time-series revenue forecasts, and determine whether estimated growth in web traffic has incremental predictive value for firm revenues over and above the best time-series forecast. Analysts' revenue forecasts are introduced in Section 4. There we explore the properties of these forecasts and analyze whether historical revenue growth and estimated growth in web traffic have incremental predictive value over them. Section 5 examines whether these growth measures can be used to predict abnormal returns for our internet stock sample. The final section contains a summary of our results, our conclusions, and a discussion of potential extensions to our analysis.

1. The Data and Descriptive Statistics

Our initial sample consists of all firms appearing on the InternetStockList (compiled by internet.com) as of January 31, 2000.[4] From this sample we retain only those firms that we judge to be primarily portals, content/community providers (collectively referred to as p/c firms below), or e-tailers.[5] To this list we add Netscape, geocities, broadcast.com, Excite, Onsale, and Xoom.com which had been public, but which were acquired by other firms prior to this date. (Adding these formerly-public firms ensures that our sample is complete as of the beginning of 2000.[6]) This leaves us with 95 firms. The appendix provides a list of these firms. Our sample period covers seven quarters, beginning with the fourth calendar quarter of 1998 (as this is the first quarter for which web traffic data is available) and ending with the second calendar quarter of 2000. Twenty-one of our firms were public during all seven quarters, while six were public during two or fewer quarters. The average number of quarters for which our firms have been publicly traded is slightly less than 5.

1.1. Financial Data

Each of our sample firm's historical net revenues is recorded, beginning with the quarter of its initial public offering, or the second calendar quarter of 1998, whichever is later, and ending with the second calendar quarter of 2000.[7] As will become clear below, we start with the second quarter of 1998 in order to have enough data to form time-series revenue forecasts for the fourth quarter of 1998 and beyond. We also collect all the one-period ahead quarterly revenue forecasts of analysts appearing on the *I/B/E/S* detailed sales forecast database for the fourth calendar quarter of 1998 and beyond. Forty-six of our sample firms have analyst revenue forecasts for at least one quarter of our sample period.

The top rows of Table 1, panels A and B, provide descriptive financial statistics for our firms. The average (median) level of quarterly revenues is only \$67.8 million (\$9.3 million) for the p/c firm subsample and \$48.0 million (\$17.1 million) for the e-tailers. In contrast, the average (median) market capitalization of these companies is quite high, at \$8.3 billion (\$645.6 million) for the p/c firms, and \$2.5 billion (\$314.5 million) for the e-tailers. The highest quarterly revenues for our sample is \$1.9 billion, for Yahoo! during the fourth quarter of 1999, while the highest market capitalization is \$190.5 billion, for America Online on June 30, 2000. Not surprisingly, the mean (median) market value/revenue ratio is a very high 135.7 (70.7) for the p/c firms and 80.2 (23.9) for the e-tailers. The maximum market value/revenue ratio is 1,243.8, for musicmaker.com at the end of the third quarter of 1999. As expected for this industry, quarter-to-quarter revenue growth is quite high, averaging 46.4 percent (with a median of 26 percent) for the p/c firms and 43.9 percent (median of 19 percent) for the e-tailers. Growth differs greatly across firms, though, ranging from a low of −48.1 percent to a high of 804 percent for the p/c firm subsample and from a low of −78.5 percent to a high of 702 percent for the e-tailers.

The untabulated Pearson correlation between current and one-quarter lagged revenue growth reveals a striking difference between the two subsamples. While the correlation for the p/c firms is a significant 24.2 percent, it is insignificant for the e-tailers. The weak e-tailer correlation is likely due, at least in part, to seasonalities which are widespread in the retail industry. This suggests that time-series revenue forecasts will be more accurate for the p/c firms than for the e-tailers and will have greater incremental predictive power for revenues above analysts' forecasts.

Table 1. Descriptive statistics for our sample firms, by firm type, 4th quarter 1998–2nd quarter 2000. The sample consists of 95 publicly traded firms listed on internet.com's InternetStockList (as of January 31st, 2000) that we classified as either portals and content/community providers (the p/c firms) or e-tailers. See the Appendix for a list of the sample firms. Market value is the market value of common shareholders' equity and is calculated using the closing price on the day after the earnings announcement, multiplied by the total number of shares outstanding at that time. Unique visitors is the estimated number of different individuals who visit a firm's web site(s) during a particular quarter. Pageviews is the number of unique visitors multiplied by both the average usage days per visitor and average daily unique pages viewed per visitor in a month. Minutes is the number of minutes spent on a firm's web site(s) and equals the number of pageviews multiplied by the average minutes spent per unique page. Growth is measured quarter-to-quarter. All numbers (except ratios and percentages) are in millions.

Variable	*N*	Mean	Median	Std Dev	Minimum	Maximum
Panel A: The p/c firms						
Revenues	215	67.8	9.3	270.1	0.1	1929.0
Market value	211	8308.8	645.6	28498.9	15.7	190505.4
Market value/Revenues	211	135.7	70.7	189.9	4.1	1183.9
Revenue growth (%)	215	46.4	26.0	92.5	−48.1	804.2
Unique visitors	195	8.4	3.5	11.4	0.2	48.4
Pageviews	195	1084.5	59.3	3409.0	1.9	22991.0
Minutes	195	1105.8	64.3	3515.0	1.7	24893.4
Unique visitor growth (%)	191	4.6	2.0	17.6	−35.4	108.2
Pageview growth (%)	191	7.5	0.0	31.7	−55.4	218.2
Minutes growth (%)	191	8.4	1.0	38.3	−63.2	253.6
Panel B: The e-tailers						
Revenues	183	48.0	17.1	95.1	0.2	676.0
Market value	178	2528.2	314.5	5854.8	10.8	29057.7
Market value/Revenues	178	80.2	23.9	159.2	0.9	1243.8
Revenue growth (%)	183	43.9	19.0	97.6	−78.5	702.3
Unique visitors	164	3.0	1.4	3.6	0.2	16.6
Pageviews	164	302.4	22.2	1220.5	0.6	9032.0
Minutes	164	306.7	18.9	1233.3	0.7	9032.0
Unique visitor growth (%)	159	7.9	5.0	24.9	−41.4	109.4
Pageview growth (%)	159	16.0	6.0	58.3	−58.7	535.0
Minutes growth (%)	158	16.0	6.0	64.8	−54.2	640.8

1.2. Web Usage Data

The web traffic data used in this study come from Media Metrix, which has the longest time series of such data of any independent internet audience measurement firm, and which was described in a *Wall Street Journal* article as the most widely used web rating company.[8] They have more than 600 clients, including financial services companies, advertising agencies, and e-commerce marketers. Media Metrix provided us with their monthly Web Report from October 1998 through September 2000.[9] This report gives a number of different metrics for each web site with a projected reach of 0.4% or higher.[10] (For a firm with more than one web site, these metrics are cumulated over all the web sites it owns.) Each month's Web Report is normally released to clients (who pay a fee to obtain access to the report) around the 20th of the following month. At the same time the company also issues a press release listing the number of unique visitors to the top 50 web sites during the month. This information, however, is a very small subset of that contained in the monthly Web

Report. Eighty-six of our sample firms have web usage data reported in at least one Web Report.

We choose to focus on three measures of internet usage: "unique visitors," "pageviews," and "minutes." The first two are among the most often-cited measures in the popular press. For a given firm, unique visitors is the number of different individuals who visit the firm's web site(s) during a particular month. The unique visitor numbers are taken directly from Media Metrix's monthly Web Report. Pageviews is the number of pages viewed by those individuals visiting the firm's web site(s) during the month. While it is not directly reported by Media Metrix (there is no universally agreed-upon definition of this measure), we calculate it by multiplying together three measures that they do provide: (1) the number of unique visitors, (2) the average usage days per visitor in the month, and (3) the average daily unique pages viewed per visitor in the month.[11] Minutes is the number of minutes spent on the firm's web site(s) during the month. It is included in our analysis because it is a measure of intensity of web site usage. While not explicitly reported by Media Metrix, we compute it by multiplying the number of pageviews by the average minutes spent per unique page during a day (the latter measure is provided by Media Metrix).

The bottom rows of Table 1, panels A and B, provide descriptive statistics on the three measures of web traffic for our sample firms. As was the case for the financial measures, the p/c firms swamp the e-tailers with respect to the magnitude of web usage. The number of unique visitors per month averages 8.4 million (the median is 3.5 million) for the p/c firms and 3.0 million (median of 1.4 million) for the e-tailers. Web site pageviews per month averages 1.1 billion (median of 59.3 million) for the p/c firms and 302 million (median of 22.2 million) for the e-tailers, while the average number of minutes spent at web sites per month is 1.1 billion (median of 64.3 million) for the p/c firms and 307 million (median of 18.9 million) for the e-tailers.

As expected, web usage is growing over time. The average quarter-to-quarter growth in unique visitors is 4.6 percent (median of 2.0 percent) for the p/c firms and 7.9 percent (median of 5.0 percent) for the e-tailers. Pageview growth averages 7.5 percent (median of 0.0 percent) for the p/c firms and 16.0 percent (median of 6.0 percent) for the e-tailers. Growth in minutes spent on the web sites averages 8.4 percent (median of 1.0 percent) for the p/c firms and 16.0 percent (median of 6.0 percent) for the e-tailers. While the web traffic measures are greater in magnitude for the p/c firms, they are growing faster for the e-tailers.

2. The Relation Between Revenues and Web Usage

In this section we examine the relation between internet firms' revenues and measures of web usage. The results of this examination will serve as a foundation for our subsequent analysis of the incremental predictive power of estimated web traffic growth for contemporaneous revenue growth, over and above time-series and analysts' forecasts. Ex-ante, we expect the revenues of the e-tailers and the p/c firms to be positively related to the three web traffic metrics we study—unique visitors, pageviews, and minutes. For the e-tailers, unique visitors reflects the number of potential customers, while pageviews and minutes capture the intensity of their shopping experience. For p/c firms, whose revenues come mainly from advertising, the booking of ad revenues is tied directly to the number of pageviews on which

Table 2. Correlation between quarter-to-quarter revenue growth and both actual and estimated quarterly growth in web usage, 4th quarter 1998–2nd quarter 2000. Quarterly web usage growth is estimated by assuming a constant change for all remaining months in the quarter. Panel A presents the correlation between revenue growth and actual growth in web usage. Panel B reports the correlation between revenue growth and estimated growth in web usage, separately for each month of the quarter.

	The p/c Firms		The e-Tailers	
Usage measure	*N*	Pearson	*N*	Pearson
Panel A: Correlation Between Revenue Growth and Actual Growth in Web Usage				
Unique visitors	170	0.212**	136	0.579**
Pageviews	170	0.109	136	0.638**
Minutes	170	0.074	136	0.634**
Panel B: Correlation Between Revenue Growth and Estimated Growth in Web Usage				
First month of quarter:				
Unique visitors	171	−0.073	140	0.124
Pageviews	171	0.057	140	0.189*
Minutes	171	0.008	139	0.152*
Second month of quarter:				
Unique visitors	171	0.172*	138	0.281**
Pageviews	171	0.078	138	0.359**
Minutes	171	0.088	137	0.309**
Third month of quarter:				
Unique visitors	170	0.234**	139	0.612**
Pageviews	170	0.092	139	0.649**
Minutes	170	0.078	137	0.654**

*(**) denotes significance at the 10% (5%) level using a two-sided test.

each advertiser's banners appear. Additionally, the more unique visitors at a p/c firm's web sites, or the more minutes the visitors spend at the sites, the greater the number of new advertisers the firm will presumably attract and the higher the rates it will likely be able to charge in the future.

2.1. *The Correlation Between Revenue Growth and Actual Web Usage Growth*

Table 2, panel A presents the correlations between actual quarter-to-quarter web usage growth and contemporaneous revenue growth; the numbers are generally supportive of the above conjectures.[12] They also highlight a striking difference between our two subsamples. For the e-tailers the Pearson correlations range from 57.9 to 63.8 percent and are all significant. For the p/c firms, in contrast, the correlations lie between 7.4 and 21.2 percent and are only significant for unique visitors. In conjunction with subsample differences in the correlation between past and current revenue growth, these results strongly suggest that web traffic measures will provide significant incremental predictive power for current revenue growth over time-series models involving past revenue growth for the e-tailers, but not necessarily for the p/c firms.

There are (at least) two possible reasons for the weaker p/c firm correlations. One is the relative diversity of the p/c firms, which range from AOL, a firm that generates a substantial amount of revenues by selling access to the web, to Yahoo! which is in large part a search engine, to Marketwatch.com, which is a site providing financial content. While for any *given* firm the correlation over time between web usage growth and revenue growth might be high, the diversity across firms might obscure the relation cross-sectionally. This problem is less likely to arise for the e-tailers, as their businesses are more homogeneous.[13] Another possible reason for weaker p/c firm correlations is that, for these firms, there may be a greater time lag between usage growth and revenue growth. It may take one or more quarters, for example, before an increase in current web traffic has an impact on the number of advertisers and advertising rates.

2.2. *The Correlation Between Revenue Growth and Estimated Web Usage Growth*

Actual traffic growth at a firm's web site(s) in a given quarter t (denoted by U_t) generally cannot be used for forecasting that quarter's revenue growth since actual usage for the quarter is usually not available before the firm's earnings (and revenues) are announced. For forecasting purposes, therefore, it is necessary to estimate the quarter's web usage growth. Since Media Metrix releases web traffic numbers monthly, this estimate will change as the quarter progresses, presumably becoming more accurate. Estimates made in the first month of the quarter (after the Web Report release date that month), for example, will have use of traffic numbers up through the last month of the prior quarter. (Recall that these are the latest figures available in the first month of the current quarter.) Similarly, estimates made in the second (third) month (again, after the Web Report release date) will have use of actual web traffic data through the first (second) month of the current quarter.

As of any particular date during quarter t, the estimated web usage for the quarter, denoted by UF_t, is computed by adding together (1) the monthly usage numbers already released by Media Metrix for the quarter (if any) and (2) forecast(s) for all remaining months in the quarter. These forecasts are derived assuming a constant change in *monthly* usage.[14] To illustrate, consider the estimation of web usage for the first quarter of 1999, computed as of March 31. At that point the January and February usage numbers will have been released. Our first quarter 1999 forecast will then be equal to the sum of the January and February numbers, plus the March forecast. That forecast is calculated by taking February usage and adding to it the change in usage between January and February.[15]

Panel B of Table 2 presents the correlations between realized quarterly revenue growth and *estimated* quarterly web usage growth, separately by month of the quarter in which the estimate is made. As was the case for actual web traffic growth, these correlations are uniformly higher for the e-tailers than for the p/c firms. Moreover, while they are significant for the e-tailers in all cases but one, for the p/c firms there is significance only for unique visitors in months two and three. Also as expected, the correlations are generally lower than those calculated using the actual web traffic numbers. Furthermore, they generally increase as the quarter progresses because more actual web usage numbers become available and can be used in the forecast. The correlations for the third month, in fact, are very similar to those attained using the actual data.

3. Time-series Revenue Forecasts and the Incremental Role of Web Traffic

The previous analysis has shown that both past revenue growth and web usage growth are correlated with current quarterly revenue growth for one or both of our subsamples (see sections 1.1 and 2, respectively). With this as a foundation, we begin this section with an examination of the predictive ability of several time-series forecasts. This analysis is expected to be especially useful in the valuation of firms with little or no analyst coverage, where historical revenues, out of necessity, take on an increased importance for forecasting purposes. This is followed by an exploration of whether web usage growth provides incremental forecasting ability for current quarterly revenues over and above the best time-series forecasting model.

3.1. The Predictive Ability of Time-series Revenue Forecasts

The first revenue forecast we consider in this subsection is simply the prior quarter's revenues, R_{t-1}. This is a random walk revenue model, and assumes no growth. This forecast will be denoted below by *PR* (for prior revenue). Our second forecast assumes that the change in revenues for the current quarter equals that of the previous quarter. Denoted by *CRC* (for constant revenue change), it is given by

$$CRC_t = R_{t-1} + (R_{t-1} - R_{t-2}) \tag{1}$$

Our final forecast is a variant on the second, and assumes that the growth rate in current-quarter revenues equals the previous quarter's growth rate. Denoted by *CRG* (for constant revenue growth), it is calculated as

$$CRG_t = R_{t-1} \cdot \frac{R_{t-1}}{R_{t-2}} \tag{2}$$

While these three time-series forecasts are intuitively appealing, there are, of course, others that can be constructed from the historical revenue data. (Several alternatives were examined, none of which demonstrated superior performance.) It must be recognized, however, that the lack of more than a few years of data for most of the sample firms precludes the development of more complex time-series forecasts, such as those incorporating seasonal patterns.

We employ three separate criteria to compare our forecasts. The first is the mean percentage error (MPE), which is the average percentage difference between realized and forecasted revenues. The second is the mean absolute percentage error (MAPE), which is the average absolute percentage difference between realized and forecasted revenues. The third is the percentage of firm-quarters in which each measure has the smallest MAPE.

Table 3 presents the forecast comparisons, separately for the p/c firms (panel A) and the e-tailers (panel B). For the p/c firms the *CRC* forecast is clearly superior along all dimensions. Its MPE error of 0.2 percent is over 14 percentage points smaller in magnitude than that of its closer rival (the *CRG* forecast), while its MAPE is better than that of the next-best forecast (the *PR* forecast) by over 8 percentage points. It also has the lowest MAPE in a plurality (41.3 percent) of the firm-quarters.

Table 3. Comparison of revenue forecasts based on historical quarterly accounting numbers, 4th quarter 1998–2nd quarter 2000. The PR forecast model is a random walk revenue model, assuming no growth. The CRC model assumes revenue change in the current quarter equal to that of the previous quarter, while the CRG model assumes a revenue growth rate in the current quarter equal to that of the previous quarter.

Forecast Model	N	Mean Percentage Error (MPE)	Mean Absolute Percentage Error (MAPE)	Percentage of Quarters with the Lowest MAPE
Panel A: The p/c firms				
Prior revenue (PR)	150	19.8	23.3	24.0
Constant revenue change (CRC)	150	0.2	15.1	41.3
Constant revenue growth (CRG)	150	−14.8	24.4	34.7
Panel B: The e-tailers				
Prior revenue (PR)	122	10.7	26.9	41.8
Constant revenue change (CRC)	122	−5.1	39.1	21.3
Constant revenue growth (CRG)	122	−53.7	75.3	36.9

While the *CRC* forecast is not as clearly superior for the e-tailers, it excels when all criteria are considered together. Its −5.1 percent MPE is over 5 percentage points smaller in magnitude than that of its closer competitor (the *PR* forecast). Its MAPE of 39.1 percent, though, is more than 12 percentage points worse than the best forecast along this dimension (again, the *PR* forecast). However, its relatively high MAPE is driven by a few very large absolute errors, as reflected by the fact that the untabulated median absolute percentage error is virtually the same for this forecast (at 17.7 percent) as it is for the *PR* forecast (at 17.5 percent). The percentage of quarters in which the *CRC* forecast has the lowest MAPE is the smallest of all the forecasts. But, this is because it often comes in second (due to a high correlation with the *CRG* forecast). In untabulated results we find the frequency that it is either first or second to be 88.5 percent, which is much higher than that of the *PR* forecast (at 48.4 percent) or that of the *CRG* forecast (at 63.1 percent).

Regardless of which time-series model is used, though, it is apparent that forecast accuracy (measured by either the MPE or the MAPE) is higher for the p/c firms than for the e-tailers. This is not surprising, given our finding of a strong (weak) correlation between current and past revenue growth for the p/c firms (e-tailers).

3.2. *The Incremental Predictive Value of Web Traffic Data Over Historical Revenues*

We now examine whether estimated web traffic growth has incremental value over the best-performing time-series model, the *CRC* forecast, in the prediction of contemporaneous quarterly revenues. We do so by regressing current quarterly revenue growth on the *CRC* forecast (divided by R_{t-1}) and estimated web usage growth (growth in unique visitors, pageviews, and minutes). Ex-ante, we expect that web usage data should have incremental predictive value, if for no other reason than that it is released monthly by Media Metrix, and so is more up-to-date than time-series forecasts (which can only be updated quarterly).

In Table 4, panel A, we first present the regression results with actual web usage growth substituted for estimated growth as an independent variable. Not surprisingly given our

Table 4. Summary of estimation results from regressing quarterly revenue growth on the CRC (constant revenue change) forecast of revenue growth and web usage growth, 4th quarter 1998–2nd quarter 2000. Growth is measured quarter-to-quarter. The following panels present the estimated coefficients on the forecasting variables, the number of observations used in the regression (N), and the adjusted R-squared (Adj. R^2). To save space, the estimated intercepts are not reported. Panel A reports the regression results based on actual web usage growth, by firm type and web usage measure. Panel B presents the findings based on estimated web usage growth, by firm type and web usage measure, separately for each month of the quarter.

	The p/c firms				The e-tailers			
Usage Measure	CRC Forecast	Actual Usage	N	Adj. R^2	CRC Forecast	Actual Usage	N	Adj. R^2
Panel A: Regression of quarterly revenue growth on the CRC forecast of revenue growth and actual web usage growth								
Unique visitors	0.972**	−0.001	143	0.076	−0.062	1.791**	114	0.448
Pageviews	1.004**	−0.089	143	0.079	0.100	0.988**	114	0.538
Minutes	0.992**	−0.089	143	0.080	0.092	0.923**	112	0.512
Panel B: Regression of quarterly revenue growth on the CRC forecast of revenue growth and estimated web usage growth								
First month of quarter:								
Unique visitors	0.951**	−0.133	144	0.071	0.093	0.137	117	−0.003
Pageviews	0.961**	−0.005	144	0.062	0.084	0.114	117	0.004
Minutes	0.955**	−0.024	144	0.063	0.088	0.082	116	−0.004
Second month of quarter:								
Unique visitors	0.960**	−0.001	144	0.062	0.037	0.403**	115	0.052
Pageviews	0.992**	−0.052	144	0.065	0.117	0.476**	115	0.202
Minutes	0.985**	−0.041	144	0.065	0.141	0.346**	114	0.132
Third month of quarter:								
Unique visitors	0.933**	0.132	143	0.065	−0.067	1.551**	117	0.439
Pageviews	1.002**	−0.087	143	0.066	0.088	0.871**	117	0.546
Minutes	0.985**	−0.071	143	0.066	0.091	0.816**	115	0.543

*(**) denotes significance at the 10% (5%) level using a two-sided test.

previous analyses, the coefficient on the *CRC* forecast is significantly positive for the p/c firms, but is insignificant for the e-tailers. In contrast, the coefficient on web usage growth for the p/c firms is insignificant, regardless of the web usage metric employed, while it is significantly positive for the e-tailers. Clearly, perfect foreknowledge of actual web traffic growth would be valuable, over and above historical revenues, in the prediction of the current quarter's revenues for the e-tailers; the same does not hold for the p/c firms.

A similar pattern is evident when estimated web traffic growth is entered into the regressions (panel B of Table 4). For each of the three months of the quarter, the coefficient on the *CRC* forecast remains significantly positive for the p/c firms, while the coefficient on estimated usage growth is uniformly insignificant over all three web usage measures. In contrast, while the coefficient on the *CRC* forecast for the e-tailers is never significant, the coefficient on estimated web usage growth is significantly positive for all three web traffic metrics in months two and three. Moreover, the coefficient's (untabulated) significance level increases over the months of the quarter. This is as expected, given that the estimates of web usage growth become more accurate as time passes during a quarter. The results of this analysis suggest that estimated web usage growth can provide predictive power for e-tailers' current quarterly revenues over and above historical revenues.[16]

4. Analysts' Revenue Forecasts and the Incremental Roles of Historical Revenues and Web Traffic Growth

The focus of our analysis now turns to a subset of our sample firms—those followed by analysts—in order to examine the predictive ability of analysts' consensus (mean) revenue forecasts. We also explore whether historical revenue growth and growth in web traffic have incremental forecasting value over and above analysts' forecasts.

As a prelude to our analysis, each firm's quarterly consensus forecast must be computed. There are many possible ways to do so. At one extreme, we could calculate just one consensus forecast (by averaging all of the forecasts made during the quarter). At the other extreme, we could continuously update the consensus throughout the quarter. Since one of our goals is to test whether estimated quarterly web usage growth provides incremental predictive value over analysts' forecasts, and these estimates are updated monthly, we choose to calculate one consensus forecast for each month of the quarter. To do so we subdivide each quarter into three intervals (months). The first month begins on the release date of the first Media Metrix Web Report for the quarter, or the time of the previous quarter's earnings announcement, whichever is later, and ends four weeks after the Web Report's release. The second (third) month of the quarter begins on the release date of the second (third) Media Metrix Web Report for the quarter and ends four weeks later.[17] If the quarter's earnings announcement does not occur before the end of the third month, however, we extend that interval until the earnings announcement date.[18] If an analyst released multiple reports during any given month, we keep only the earliest one for the purposes of computing the consensus forecast. We do so in order to avoid double-counting as well as to minimize any potential time advantage the analyst might have over the historical revenue and web usage data. The remaining forecasts released during the month are then averaged to compute the consensus.[19]

4.1. *The Predictive Ability of Analysts' Consensus Revenue Forecasts*

Not surprisingly, the (untabulated) Pearson correlation between current quarterly revenue growth and analysts' revenue forecasts (divided by R_{t-1}) is quite high. For the p/c firms the correlation is 71.8 percent, while for the e-tailers it is an even higher 96.0 percent. More interesting are the descriptive statistics pertaining to the percentage consensus analyst forecast errors for the two subsamples. For each month m of quarter t, we calculate this error as

$$\frac{R_t - AF_{m,t}}{R_t}$$

where $AF_{m,t}$ is the consensus forecast for that month. As reported in Table 5, panels A and B, the average consensus forecast errors are significantly positive in each individual

Table 5. Descriptive statistics on percentage analyst consensus revenue forecast errors, by firm type, for 386 firm-month observations, 4th quarter 1998–2nd quarter 2000. The consensus forecasts are computed by averaging all analyst forecasts made during a given month. If an analyst released multiple reports during the month, only the earliest forecast is used in the computation. The first month of the quarter begins on the release date of the first Media Metrix Web Report for the quarter, or the announcement date of last quarter's earnings, whichever is later, and ends 4 weeks after the Media Metrix release date. The second month of the quarter begins on the release date of the second Media Metrix Web Report for the quarter and ends 4 weeks later. The third month of the quarter begins on the release date of the third Media Metrix Web Report for the quarter and ends either 4 weeks later or at the earnings announcement date for the quarter, whichever is later.

		Calendar Quarter						
	All	1998 Q4	1999 Q1	1999 Q2	1999 Q3	1999 Q4	2000 Q1	2000 Q2
Panel A: The p/c firms								
N	219	10	18	22	35	45	50	39
Mean	−11.9	−9.5	−10.6	−12.4	−15.4	−13.3	−13.1	−6.7
Std. Dev.	14.3	13.1	10.5	6.0	13.8	12.2	19.7	12.9
Median	−11.5	−6.8	−8.2	−11.7	−13.8	−12.9	−13.7	−5.8
% > 0	91.8	90.0	89.0	100.0	91.0	96.0	94.0	82.0
MAPE	14.6	10.0	12.0	12.0	16.0	15.0	19.0	10.0
Prob > $\|T\|$	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00
Prob > $\|W\|$	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Panel B: The e-tailers								
N	167	9	11	19	33	31	38	26
Mean	−6.0	−2.1	−11.9	−12.9	−6.6	−2.0	−11.9	5.1
Std. Dev.	21.2	28.2	8.1	15.5	24.2	31.1	10.0	15.4
Median	−9.5	−13.2	−9.8	−12.1	−10.2	−9.9	−13.9	1.0
% > 0	78.6	78.0	100.0	79.0	88.0	84.0	87.0	38.0
MAPE	15.5	22.0	12.0	15.0	16.0	21.0	13.0	11.0
Prob > $\|T\|$	0.00	0.83	0.00	0.00	0.13	0.72	0.00	0.11
Prob > $\|W\|$	0.00	0.57	0.00	0.00	0.00	0.02	0.00	0.22

MAPE denotes mean absolute percentage error.
Prob > $|T|$ denotes the p-value of the t-statistic for the hypothesis that the population mean is zero.
Prob > $|W|$ denotes the p-value of the Wilcoxon signed rank statistic for testing the hypothesis that the population median is zero.

quarter over our sample period, with the one exception of the e-tailers in the second quarter of 2000.[20] For the p/c firms the average consensus forecast error over all quarters is 11.9 percent, ranging from a low of 6.7 percent during the second quarter of 2000 to a high of 15.4 percent in 1999's third quarter. The mean absolute percentage error (MAPE), meanwhile, is 14.6 percent, and ranges from 10 percent in both the fourth quarter of 1998 and second quarter of 2000 to 19 percent during the first quarter of 2000. Moreover, the vast majority (between 82 and 100 percent) of consensus analyst forecast errors are positive. For the e-tailers the average consensus forecast error over all quarters is 6.0 percent. Leaving aside the second quarter of 2000, the average error varies between 2.0 percent in the fourth quarter of 1999 to 12.9 percent in 1999's second quarter. The average percentage error in the second quarter of 2000 is −5.1 percent.[21] Over all quarters, the average MAPE is 15.5 percent, ranging from 11 percent in 2000's second quarter to 22 percent in the fourth quarter of 1998. Aside from the second quarter of 2000, the large majority (between 78 and 100 percent) of consensus forecast errors are positive. In 2000's second quarter the percentage drops to 38 percent.[22]

That analysts have been *underestimating* revenues on a regular basis stands in contrast to the preponderance of empirical research from other industries which finds an *optimistic* bias to analysts' earnings forecasts and which shows that only about 50 percent of the consensus forecasts are pessimistic.[23] There are several possible explanations for the observed underestimation in analysts' revenue forecasts. One possibility is that internet firms are consistently surprising analysts by their strong revenue growth. If this is the case, then we should expect to see a decrease in the observed pessimism over time. It is true that in the most recent quarter (the second quarter of 2000) analysts did not exhibit any pessimism with respect to the e-tailers and the level of underestimation for the p/c firms was the smallest of all the quarters. However, the lack of any discernible decrease in pessimism over the remaining six quarters gives rise to doubts that this is the entire explanation. Another possibility is that these firms manage their revenues upward, so as to beat analysts' forecasts.[24] A problem with this explanation, though, is that it does not allow for analysts to adjust their forecasts to take into account such revenue manipulation. A third possibility is that analysts deliberately bias their forecasts downward so that the firms they cover can report greater than expected revenues and, as a result, boost their stock prices. A final possibility is that analysts are simply not able to properly interpret the available data for these high growth firms, as prior literature has shown to be the case with respect to analysts' earnings forecasts in other high growth firms.[25] We leave a further discussion and analysis of these possibilities to future work.

4.2. The Incremental Predictive Value of Historical Revenues and Web Usage Data Over Analysts' Forecasts

We turn first in this subsection to an examination of whether historical revenue numbers have incremental predictive value for current quarterly revenues (R_t) over analysts' consensus forecasts. To do so we regress current quarterly revenue growth on the analysts' consensus revenue forecast (divided by R_{t-1}) and the best time-series forecast, *CRC* (again divided by R_{t-1}). Table 6, panel A, presents the regression results for each of our two subsamples. Across all the months of a quarter, and for both the p/c firms and the e-tailers, the coefficient on the analysts' revenue forecast is significantly positive. This is not surprising, given the strong correlation previously found between analysts' forecasts and realized quarterly revenues. The *CRC* time-series forecast adds significant incremental predictive value for p/c firms' revenues in months one and two of the quarter, as reflected by its significantly positive coefficient for those months. The incremental predictive value of the *CRC* forecast is much weaker for the e-tailers, though, as the time-series coefficient exhibits (marginal) significance only in month one. The untabulated t-statistic on the time-series forecast regression coefficient decreases for both subsamples as we move from month one to month three. This is reasonable, given that analysts' forecasts are expected to become relatively more accurate as the quarter progresses. The incremental value of the time-series forecast (especially for the p/c firms) suggests that either analysts are not appropriately taking into account the autocorrelation structure of revenue growth or they are intentionally introducing bias into their forecasts.

We next consider whether estimated quarterly web traffic growth has incremental predictive value for revenues over and above analysts' forecasts, by regressing current quarterly revenue growth on the analysts' consensus revenue forecast (divided by R_{t-1}) and estimated web traffic growth. Since a necessary condition for estimated web traffic growth to have incremental predictive ability is that *actual* web traffic growth have this power, we first present, in panel B of Table 6, the results of substituting actual web usage growth for estimated growth in the regression equation.

As before, the analysts' revenue forecast is significantly associated with realized quarterly revenues for each month of the quarter and for both the p/c firms and the e-tailers. As reflected in its regression coefficient, actual web traffic growth has significant incremental forecasting value for the p/c firms in month three, across all three metrics, and in month one for the

Table 6. Summary of estimation results from regressing quarterly revenue growth on the analysts' forecast of revenue growth and either the CRC (constant revenue change) forecast of revenue growth or web usage growth (or both), 4th quarter 1998–2nd quarter 2000. Growth is measured quarter-to-quarter. The following panels present the estimated coefficients on the forecasting variables, the number of observations used in the regression (N), and the adjusted R-squared (Adj. R^2). To save space, the estimated intercepts are not reported. Panel A presents the estimation results from regressing quarterly revenue growth on the analyst and CRC forecasts. Panel B (Panel C) reports the results from regressing revenue growth on the analysts' forecast and actual (estimated) web usage growth. Panel D (Panel E) presents the results from regressing quarterly revenue growth on the analyst and CRC forecasts, and actual (estimated) web usage growth. All results are presented by firm type and month of the quarter, and separately for each web usage measure (when they are included in the regressions).

	The p/c firms				The e-tailers			
	Analyst Forecast	CRC Forecast	N	Adj. R^2	Analyst Forecast	CRC Forecast	N	Adj. R^2
Panel A: Regression of quarterly revenue growth on the analysts' forecast of revenue growth and CRC forecast of revenue growth								
First month of quarter	0.675**	0.439**	99	0.399	1.285**	0.074*	71	0.937
Second month of quarter	0.910**	0.279**	55	0.611	1.262**	0.194	32	0.918
Third month of quarter	1.468**	0.283	35	0.726	0.926**	−0.088	27	0.795
Panel B: Regression of quarterly revenue growth on the analysts' forecast of revenue growth and actual web usage growth								
First month of quarter:								
Unique visitors	0.698**	0.311**	98	0.404	1.198**	0.248**	76	0.948
Pageviews	0.747**	0.093	98	0.370	1.197**	0.126**	76	0.948
Minutes	0.738**	0.079	98	0.368	1.181**	0.144**	75	0.950
Second month of quarter:								
Unique visitors	0.785**	0.218	51	0.495	1.055**	0.570**	36	0.949
Pageviews	0.820**	−0.015	51	0.466	1.080**	0.250**	36	0.949
Minutes	0.818**	−0.019	51	0.467	1.072**	0.248**	36	0.950
Third month of quarter:								
Unique visitors	1.274**	0.247**	33	0.855	0.894**	0.277	26	0.802
Pageviews	1.311**	0.220**	33	0.868	0.873**	0.210**	26	0.831
Minutes	1.265**	0.222**	33	0.865	0.901**	0.182**	26	0.825

(*continued*)

Table 6. (*continued*).

Panel C: Regression of quarterly revenue growth on the analysts' forecast of revenue growth and estimated web usage growth

	The p/c firms				The e-tailers			
Usage measure	Analyst Forecast	Estimated Usage	N	Adj. R^2	Analyst Forecast	Estimated Usage	N	Adj. R^2
First month of quarter:								
Unique visitors	0.762**	−0.053	98	0.364	1.274**	0.035	76	0.945
Pageviews	0.742**	0.015	98	0.360	1.279**	−0.010	76	0.945
Minutes	0.747**	0.000	98	0.359	1.277**	0.008	76	0.945
Second month of quarter:								
Unique visitors	0.778**	0.153**	51	0.510	1.207**	0.119	36	0.937
Pageviews	0.816**	0.016	51	0.468	1.208**	0.078	36	0.938
Minutes	0.817**	0.006	51	0.466	1.216**	0.065	36	0.937
Third month of quarter:								
Unique visitors	1.303**	0.174*	33	0.847	0.915**	0.085	28	0.781
Pageviews	1.321**	0.171*	33	0.847	0.884**	0.148	28	0.799
Minutes	1.286**	0.118	33	0.839	0.902**	0.140	28	0.799

Panel D: Regression of quarterly revenue growth on the analysts' forecast of revenue growth, CRC forecast of revenue growth, and actual web usage growth

	The p/c firms					The e-tailers				
Usage measure	Analyst Forecast	CRC Forecast	Actual Usage	N	Adj. R^2	Analyst Forecast	CRC Forecast	Actual Usage	N	Adj. R^2
First month of quarter:										
Unique visitors	0.547**	0.226	0.156	87	0.333	1.165**	−0.009	0.322**	65	0.958
Pageviews	0.547**	0.275*	0.042	87	0.320	1.142**	0.004	0.173**	65	0.959
Minutes	0.543**	0.291**	0.005	87	0.316	1.098**	0.004	0.211**	64	0.962
Second month of quarter:										
Unique visitors	0.697**	0.160	0.120	46	0.550	0.912**	0.350	0.843**	30	0.942
Pageviews	0.700**	0.242*	−0.041	46	0.549	0.819**	0.294	0.486**	30	0.950
Minutes	0.694**	0.243*	−0.041	46	0.551	0.777**	0.258	0.501**	30	0.954
Third month of quarter:										
Unique visitors	1.185**	0.121	0.201	31	0.704	0.875**	−0.188*	0.363*	23	0.844
Pageviews	1.189**	0.191	0.197**	31	0.734	0.853**	−0.075	0.231**	23	0.861
Minutes	1.187**	0.217	0.201**	31	0.729	0.886**	−0.119	0.203**	23	0.860

(*continued*)

Table 6. (continued).

Panel E: Regression of quarterly revenue growth on the analysts' forecast of revenue growth, CRC forecast of revenue growth, and estimated web usage growth

	The p/c firms					The e-tailers				
Usage measure	Analyst Forecast	CRC Forecast	Estimated Usage	N	Adj. R^2	Analyst Forecast	CRC Forecast	Estimated Usage	N	Adj. R^2
First month of quarter:										
Unique visitors	0.550**	0.298**	−0.057	87	0.327	1.297**	0.011	0.012	65	0.954
Pageviews	0.537**	0.295**	0.021	87	0.322	1.298**	0.012	0.001	65	0.954
Minutes	0.542**	0.294**	0.001	87	0.316	1.298**	0.011	0.003	65	0.954
Second month of quarter:										
Unique visitors	0.689**	0.172	0.094	46	0.562	1.198**	0.260	0.162	30	0.918
Pageviews	0.703**	0.221*	−0.006	46	0.541	1.138**	0.371	0.171*	30	0.924
Minutes	0.702**	0.224*	−0.009	46	0.542	1.158**	0.358	0.148	30	0.922
Third month of quarter:										
Unique visitors	1.179**	0.160	0.138	31	0.695	0.894**	−0.130	0.177	25	0.798
Pageviews	1.172**	0.216	0.149	31	0.701	0.865**	−0.079	0.174*	25	0.817
Minutes	1.174**	0.263	0.104	31	0.690	0.883**	−0.078	0.176*	25	0.820

*(**) denotes significance at the 10% (5%) level using a two-sided test.

unique visitor metric. With just one exception, actual web traffic growth also exhibits significant incremental forecasting value for the e-tailers across all months and all web traffic measures. These results indicate that perfect foreknowledge of actual quarterly web traffic growth would have significant value in predicting current quarterly revenues, over and above that of analysts' forecasts.

The results of the regressions involving the *estimates* of web usage growth are reported in panel C of Table 6. As in the previous set of regressions, the analysts' consensus forecast is significantly associated with realized quarterly revenues. Coefficient magnitudes are very similar to those previously reported. The incremental forecasting value of estimated web traffic growth is weaker than that of actual web traffic growth, though, as reflected in the generally smaller regression coefficients and (untabulated) t-statistics. For the p/c firms the only significant coefficients are those for unique visitors in months two and three and for pageviews in month two. For the e-tailers, the coefficients are uniformly insignificant. The overall weaker results for the estimated web usage numbers point to the potential value for forecasting purposes of both improving upon the web usage estimates and obtaining more timely reports of actual web traffic.

Table 6, panel D (panel E) presents the results of including both the *CRC* forecast and actual (estimated) web traffic growth along with analysts' consensus forecasts in a regression with current quarterly revenue growth as the dependent variable. As before, the coefficient on the analysts' consensus forecast is uniformly significantly positive. For the p/c firms, the coefficients on the *CRC* forecast and on actual web traffic growth are less often significant than they had been when included separately with the analysts' forecasts, while estimated web traffic growth is now insignificant in all cases. (That the coefficients on actual and estimated unique visitor growth are uniformly insignificant is likely due to the multicollinearity which exists between these measures and the *CRC* forecast.) In contrast, actual growth in web usage remains uniformly significant for the e-tailers. As was the case when usage growth was included separately with analysts' consensus forecasts, the coefficient on estimated web usage growth is smaller in magnitude than the corresponding coefficient on actual growth. Now, however, there is marginal significance for the coefficient on pageviews in months two and three and for that on minutes in month three.

Taking the results of this section's analysis as a whole, we conclude that historical revenue growth plays a greater incremental role in predicting p/c firms' revenues than in predicting the revenues of e-tailers. On the other hand, web usage growth appears to be incrementally more important for the forecasting of the revenues of the e-tailers than for predicting p/c firms' revenues. Not surprisingly, this conclusion is also consistent with the results of section 3, where it is shown that the coefficient on the *CRC* forecast is significant for the p/c firms (but not the e-tailers), while the coefficient on estimated web usage growth is significant for the e-tailers (but not the p/c firms).

5. Abnormal Returns to Portfolios Formed on the Basis of Historical Revenue Growth and Estimated Growth in Web Usage

That the p/c firms' *CRC* time-series forecasts provide incremental predictive power for current quarterly revenues over and above analysts' forecasts suggests that the firms whose *CRC* forecasts are high might earn greater abnormal returns during our sample period than those whose forecasts are low. Similarly, the finding that estimated web usage growth has incremental predictive ability for e-tailers' revenues above historical revenues suggests that the e-tailers with high estimated web traffic growth might outperform those whose estimated web traffic growth is low. While finding such differential abnormal returns would be consistent with our prior results, it would not necessarily imply that investors could have designed profitable trading strategies. To do so would require that they have advance knowledge of the relations we have found for our sample period. It should also be recognized that any documented outperformance during this period of time does not necessarily imply that abnormal profits can be earned in future periods.

To begin the analysis, we rank the p/c firms each quarter according to their *CRC* forecast of quarterly revenues (divided by R_{t-1}). The top half of the firms in each of the quarters are then placed in one portfolio and the bottom half in another. Similarly, we rank the e-tailers each quarter according to estimated web usage growth. The top-half of the firms for each of the quarters go into one portfolio and the bottom half in another.[26] For completeness, we separately rank the p/c firms (e-tailers) according to estimated web traffic growth (*CRC* forecast) and again partition the firms in each subsample into two portfolios.

For each portfolio an average buy-and-hold abnormal return (BHAR) is then calculated for the period beginning 40 trading days prior to the earnings announcement date ($t = -40$) and ending on day t, where $t \in [-39, +2]$.[27] To calculate a portfolio's average BHAR for a given time period we first compute the buy-and-hold raw return for each individual

Table 7. Average buy-and-hold abnormal returns (BHAR's), 4th quarter 1998–2nd quarter 2000. Panel A reports the BHAR's to portfolios of firms with high and low CRC (constant revenue change) forecast of revenue growth, while Panel B presents the BHAR's to portfolios of firms with high and low estimated unique visitor growth, by firm type. Portfolio BHAR differences are also presented. Reported return intervals begin 40 trading days before the earnings announcement date ($t = -40$) and end on date t, where $t \in \{-30, -20, -10, 0, +2\}$. Abnormal returns are calculated using the Amex Interactive Week Internet Index as a benchmark.

	The p/c firms			The e-tailers		
Interval	High	Low	Difference	High	Low	Difference
Panel A: Buy-and-hold abnormal returns to portfolios of firms with high and low CRC forecast of revenue growth						
[−40, −30]	−0.024	−0.010	−0.013	0.115**	−0.045	−0.069*
[−40, −20]	−0.000	−0.074**	0.073	−0.160**	−0.077*	−0.083*
[−40, −10]	0.015	−0.071**	0.086	−0.154**	−0.068	−0.085
[−40, 0]	0.042	−0.103**	0.145**	−0.152**	−0.039	−0.114
[−40, +2]	−0.001	−0.115**	0.114*	−0.182**	−0.079	−0.103
Panel B: Buy-and-hold abnormal returns to portfolios of firms with high and low estimated unique visitor growth						
[−40, −30]	−0.012	−0.025	0.013	−0.074	−0.090	0.016
[−40, −20]	0.003	−0.055	0.058	−0.145	−0.126	−0.019
[−40, −10]	0.046	−0.077	0.122**	−0.143	−0.121	−0.022
[−40, 0]	−0.008	−0.068	0.076	−0.148	−0.124	−0.024
[−40, +2]	−0.025	−0.079	0.054	−0.195	−0.151	−0.045

*(**) denotes significance at the 10% (5%) level using a two-sided test.

stock in that portfolio. From that return we subtract the buy-and-hold return on the Amex Interactive Week Internet Index for the same period, which yields the stock's abnormal return.[28] Equally weighting these individual abnormal returns gives the portfolio's average BHAR for the period.

As shown in Table 7, panel A, the p/c firms with a high *CRC* forecast of quarterly revenue growth significantly outperform those with low forecasted growth, both for the period from $t = -40$ to $t = 0$ and for the period from $t = -40$ to $t = +2$. For the e-tailers there is no period for which the return to the high growth portfolio significantly outperforms that of the low growth portfolio. This result is not surprising, given that the *CRC* forecast was not found to provide incremental predictive power for these firms' revenues.

Panel B presents the return results when the firms are partitioned based on estimated unique visitor growth. (The results for pageviews and minutes are similar and are not presented.) With the exception of the period from $t = -40$ to $t = -10$ for the p/c firms, the portfolio of firms with high estimated unique visitor growth does not perform significantly better than the portfolio of firms with low estimated growth. These generally weak findings are as expected for the p/c firms given the lack of significance of estimated web traffic growth when included in regressions alongside the CRC forecast, and the mixed results when included alongside analysts' forecasts. For the e-tailers the finding is somewhat surprising, given the incremental predictive ability of estimated web usage growth over historical revenues. However, it is consistent with the largely insignificant predictive ability of estimated growth over analysts' forecasts, to the extent that analysts' forecasts reflect the market's expectations.

6. Summary, Conclusions, and Possible Extensions

In light of the importance of revenues in the valuation of internet stocks, this paper examines the roles played by analysts, past revenues, and web usage data in the forecasting of future revenues. We begin by showing that current and past quarterly revenue growth are significantly correlated for the portal and content/community providers (the p/c firms), but not for the e-tailers. A significant relation is also found between current revenue growth and the growth in web traffic (as measured by unique visitors, pageviews, and minutes spent at an internet firm's web sites), more so for the e-tailers than for the p/c firms.

We next compare several time-series forecasting models, each based on historical revenues. As we point out, this analysis is of particular use when valuing firms with little or no analyst coverage, where historical revenue numbers take on increasing importance for forecasting purposes. We find that a forecast assuming constant quarterly revenue change performs better overall (using criteria such as average and mean absolute percentage error) than other time-series models considered. Taking this time-series forecast as a benchmark, we then examine whether estimates of web traffic growth have incremental predictive power for quarterly revenues. Here, too, our p/c firm and e-tailer subsample results differ. For the e-tailers, we do find significant incremental predictive value over and above the time-series forecast. This is not the case for the p/c firms. These findings are consistent with the weaker correlations we document between traffic growth and current quarterly revenue growth for the p/c firms relative to the e-tailers.

The analysis continues by exploring the properties of analysts' quarterly revenue forecasts. As expected, analysts' forecasts are highly correlated with realized revenues. In contrast to recent research in other industries, though, which demonstrates an *optimistic* mean bias to analysts' earnings forecasts and which finds that only about half of the consensus forecasts are pessimistic, we document that analysts almost uniformly underestimate internet firms' revenues. Among possible reasons for the observed pessimism is an inability on the part of analysts to appropriately incorporate available data for these fast growing firms into their forecasts or a desire to deliberately bias their forecasts downward so as to allow the firms they cover to report positive revenue surprises.

We then examine the extent to which historical revenue numbers and estimates of quarterly web usage growth have incremental predictive value for quarterly revenues over analysts' forecasts. The results of this analysis are mixed. We find that past revenue growth has significant incremental predictive power for the p/c firms, but not for the e-tailers. Estimated web usage growth, however, is generally not found to provide significant incremental explanatory power for either set of firms. We do document, though, that perfect foreknowledge of *actual* quarterly web usage growth would have significant predictive power over analysts' forecasts. This suggests the potential value for forecasting purposes of both improving upon the estimates of web traffic growth and obtaining more timely web usage numbers.

In future work we plan to examine in more detail the documented pessimistic bias in analysts' revenue forecasts. Among the questions to be addressed are whether the bias exists in other industries and whether it is smaller for older firms (for which there is a longer time-series of historical revenues). We also plan to explore the extent to which the forecasts of underwriting firms' analysts differ from those of other analysts. The results of these analyses will not only give us a better grasp of the factors which affect the magnitude of the observed pessimistic bias, but will also help us to understand its causes.

Appendix: The Sample Firms

This table lists the names and ticker symbols of the 95 internet companies in our sample. The sample consists of all firms which appeared on the InternetStockList, complied by internet.com, as of January 31, 2000, that we judged to be primarily portals, content/community providers, or e-tailers, as well as Netscape, geocities, broadcast.com, Excite, Onsale, and Xoom.com, which had been public, but which were acquired before this time.

	Firm Name (Previous Name)	Ticker		Firm Name (Previous Name)	Ticker
1	About.com (Miningco.com)	BOUT	34	Expedia	EXPE
2	Alloy Online Inc	ALOY	35	Fashionmall.com	FASH
3	Amazon.com	AMZN	36	FatBrain.com (Computer Literacy)	FATB
4	America Online	AOL	37	FTD.com	EFTD
5	Ashford.com	ASFD	38	Garden.com	GDEN
6	Ask Jeeves	ASKJ	39	geocities	GCTY
7	Audiohighway	AHWY	40	Go2Net	GNET
8	Autobytel.com	ABTL	41	GoTo.com	GOTO
9	Autoweb.com	AWEB	42	Health Central.com	HCEN
10	barnesandnoble.com	BNBN	43	Homestore.com	HOMS
11	Beyond.com (Software.net)	BYND	44	Hoovers	HOOV
12	Bluefly	BFLY	45	Infonautics	INFO
13	Bigstar Entertainment	BGST	46	Infoseek	SEEK
14	broadcast.com	BCST	47	Infospace.com	INSP
15	C/Net	CNET	48	Insweb	INSW
16	CareerBuilder	CBDR	49	Intelligent Life	ILIF
17	CDNow	CDNW	50	Internet.com	INTM
18	Cheap Tickets	CTIX	51	IVillage	IVIL
19	China.com	CHINA	52	Iturf	TURF
20	Comps.com	CDOT	53	Knot	KNOT
21	Crosswalk.com (Didax)	AMEN	54	Launch Media	LAUN
22	Cyberian Outpost	COOL	55	Liquid Audio	LQID
23	Cybershop	CYSP	56	Looksmart	LOOK
24	Drkoop.com	KOOP	57	Lycos	LCOS
25	Drugstore.com	DSCM	58	MapQuest.com	MQST
26	EarthWeb	EWBX	59	MarketWatch.com	MKTW
27	Ebay	EBAY	60	Medscape	MSCP
28	Edgar Online	EDGR	61	MP3.com	MPPP
29	Egghead.com	EGGS	62	Multex.com	MLTX
30	emusic.com (Goodnoise)	EMUS	63	Musicmaker.com	HITS
31	E-stamp	ESTM	64	Netscape	NSCP
32	Etoys	ETYS	65	Netradio	NETR
33	Excite	XCIT	66	Netivation.com	NTVN

(*continued*)

Appendix: The Sample Firms (continued)

Firm Name (Previous Name)	Ticker	Firm Name (Previous Name)	Ticker
67 Onsale	ONSL	82 TheStreet.com	TSCM
68 Peapod	PPOD	83 Ticketmaster Online-City Search	TMCS
69 Planetrx.com	PLRX	84 Tickets.com	TIXX
70 Preview Travel	PTVL	85 US Search.com	SRCH
71 priceline.com	PCLN	86 Ubid	UBID
72 Quepasa	PASA	87 Value America	VUSA
73 RealNetworks	RNWK	88 Verticalnet	VERT
74 Salon.com	SALN	89 Visual Data	VDAT
75 Smarterkids.com	SKDS	90 Vitaminshoppe.com	VSHP
76 SportsLine USA	SPLN	91 Walt Disney Co (The) Go.com	GO
77 Stamps.com	STMP	92 Women.com Networks	WOMN
78 Starmedia Network	STRM	93 Xoom.com	XMCM
79 Student Advantage	STAD	94 Yahoo!	YHOO
80 theglobe.com	TGLO	95 Ziff-Davis Inc—ZDnet	ZDZ
81 Talk City	TCTY		

Acknowledgments

We gratefully acknowledge the comments of the editor, Charles Lee, two anonymous referees, Jennifer Altamuro, Brad Barber, Orie Barron, Anne Beatty, Kevin Cope, Paul Fischer, Steven Huddart, Stephen Penman, Siew Hong Teoh, and seminar participants at the Pricing Internet Stocks conference of the Berkeley Program in Finance/Center for Financial Reporting and Management, the 2000 *Review of Accounting Studies* conference, Mellon Capital Management, and Penn State University. We thank Media Metrix Inc. for allowing us access to their Web Reports, and the Center for Financial Reporting and Management for providing financial support. We also gratefully acknowledge the contribution of I/B/E/S International Inc. for providing sales forecast data. B. Baik, G. Jiang, D. Li, M. Luo, U. Neubauer, A. Sribunnak, Y. Zang, and B. Zeng provided able research assistance.

Notes

1. The importance of revenues in firm valuation is emphasized in "Earth to Dot-Com Accountants," by Catherine Yang (*Business Week*, April 3, 2000, p. 41), "Pricing IPOs: Science or Science Fiction," by Ed McCarthy (*Journal of Accountancy*, September 1999, p. 53), and "No Earnings? No Problem! Price-Sales Ratio Use Rises," by Robert McGough (*Wall Street Journal*, Nov. 26, 1999, C1).
2. Academic research on internet stock valuation includes Demers and Lev (2000), Hand (2000a, 2000b), Rajgopal, Kotha, and Venkatchalam (2000), and Trueman, Wong, and Zhang (2000). All of these studies show a significant relation between either revenues or gross profits and internet stock prices.
3. We do test for stability of our findings within-sample by partitioning our time period into two subperiods—the fourth quarter of 1998 through the fourth quarter of 1999, and the first and second quarters of 2000. Untabulated regression results for these two time periods are similar, in that the significant (insignificant) coefficients in one period are usually significant (insignificant) in the other period.

4. According to internet.com, the InternetStockList is "[a] comprehensive list of the more than one hundred publicly-traded companies involved solely in Internet-related business". Firms are added to the list just after their initial public offering, and are deleted when they are no longer publicly traded. As far as we are aware, no other criteria are used to determine whether a stock is included in or deleted from the InternetStockList.
5. In classifying firms we rely primarily on the self-descriptions contained in their earnings announcements.
6. Aside from these well-known, formerly public companies, only two other firms, N2K and Ozemail, are on either the 1998 or 1999 Compustat research tapes and have a business description which would result in their being classified as either p/c firms or e-tailers. We choose not to include them in our analysis since their sales figures were only available for one quarter during that period (the fourth quarter of 1998) and because our data source does not provide any analyst forecasts for that quarter.
7. A few of our firms have fiscal quarters that do not correspond to calender quarters. For these firms we collect revenue data beginning with the fiscal quarter whose end-date falls within the second calender quarter of 1998 and ending with the fiscal quarter whose end-date falls within the second calender quarter of 2000.
8. See "The Tricky Task of Tracking Web Users" (November 22, 1999, p. C1), by Nick Wingfield.
9. An official at Media Metrix informed us that the web usage data for months prior to October 1998 is not strictly comparable to that for the post-October period due to the company's merger with ReleventKnowledge, another web rating firm, around that time.
10. Media Metrix defines *reach* as the "percentage of projected individuals ... that accessed the web content of a specific site or category among the total number of projected individuals using the web during the month." We were informed by a Media Metrix official that a typical monthly Web Report contains data for approximately 22,000 sites, operated by either privately held or publicly traded companies.
11. Media Metrix gives the precise definition of *unique visitor* as "[t]he estimated number of different individuals within a designated demographic or market break category that accessed the Web content of a specific site or category among the total number of projected individuals using the web during the month." *Average usage days per visitor* is defined by them as "[t]he average number of different days in the month, per person, in which a site or category was visited." *Average (daily) unique pages per visitor in a month* is defined as "the average number of different page requests made per day over the course of the month by those persons visiting the specific site or category."
12. In contrast to our results, Rajgopal, Kotha, and Venkatachalam (2000) do not find a significant relation between percentage change in revenues and percentage change in reach. Their analysis uses a different source for web traffic data (PC Data) than we do and is restricted to one quarter, the second quarter of 1999, in contrast to our seven-quarter time period.
13. One notable exception is eBay, which does not sell goods or services to consumers; rather, it facilitates the exchange of goods and services between individuals.
14. We assume a constant change in usage, rather than a constant growth rate in usage, because of our belief that very high web traffic growth rates in a firm's early years are not likely to be sustainable. We do not make any seasonality adjustments because a firm's usage data is available for at most seven quarters.
15. As another example, consider the estimation of web usage for the first quarter of 1999, calculated as of January 31. At that point, no actual web usage data is available for the quarter. January's web traffic, UF_J, is then estimated as $UF_J = U_D + (U_D - U_N)$, where $U_D(U_N)$ denotes web usage for December (November). The estimate of February's web traffic, UF_F, is given by $UF_F = U_D + 2 \cdot (U_D - U_N)$, while that of March, UF_M, is given by $UF_M = U_D + 3 \cdot (U_D - U_N)$. Adding together UF_J, UF_F, and UF_M, gives an estimate of web usage for the quarter of $3 \cdot U_D + 6 \cdot (U_D - U_N)$.
16. In contrast to our findings, Ittner and Larcker (1998), in a study of 73 retail branch banks of a large financial institution, show that the growth in a customer satisfaction index (another non-financial measure) has no incremental predictive power for current revenue growth over past revenue growth.
17. Alternatively, we could have chosen a shorter period following the Web Report's release over which to calculate the consensus. Doing so, however, would have resulted in fewer observations, as we would have been required to drop those firm-months in which no analyst forecasts were released during the specified time period. Note that the longer time period we use biases against our finding significant incremental predictive power for past revenues and estimated web usage growth over analysts' forecasts.
18. To illustrate the construction of these monthly intervals, consider the first quarter of 1999. The first Web Report for that quarter (the December report) was issued on January 21. Therefore, the first four-week period extends from January 21, or the date of the prior quarter's earnings announcement, if later, to February 17. The second period extends from February 22 (the release date of the January Web Report) to March 21. The

third four-week period begins on March 22 (the release date of the February Web Report) and ends on April 18, or the date of the firm's earnings report for the quarter, whichever is later.

19. By construction, then, all analysts releasing forecasts during a particular month could have used the latest Web Report in their calculations. (Although access to the Web Reports requires payment of a fee, we find that analysts' reports typically make reference to Media Metrix data, suggesting that they do have access to it.) If they were to appropriately use this data, then we would not expect to find incremental predictive value in our web usage estimates.
20. The significance level for the overall sample must be interpreted with caution, however, since the forecast errors are correlated across the months of a given quarter, and may also be correlated across firms within a month or quarter.
21. An interesting question is whether the sharp decline in internet stock prices in 2000 affected analysts' forecasting incentives and contributed to the reversal in sign of their forecast errors. If it has, then we would expect to continue to observe analyst optimism (or at least diminished pessimism) during the third and fourth quarters of 2000.
22. To check the robustness of our findings we redefined the consensus analyst revenue forecast as the average of the three most recent one-period ahead forecasts within a given quarter (consistent with the manner in which many prior studies calculate the consensus), and once again computed the average forecast error by quarter and firm type. Results obtained are very similar to those reported in the text.
23. See Abarbanell and Lehavy (2000), Brown (1999), Kasznik and McNichols (1999), Matsumoto (1999), and Richardson, Teoh, and Wysocki (1999).
24. Possible revenue management by internet firms has been a topic of discussion in many recent newspaper and magazine articles. See, for example, "Presto Chango! Sales are Huge!," by Jeremy Kahn (*Fortune*, March 20, 2000, pp. 90–96) and "Plump from Web sales, some Dot-Coms Face Crash Diet of Restriction on Booking Revenue," by Elizabeth Macdonald (*Wall Street Journal*, February 28, 2000, p. C2). It should be noted that the introduction of the Security and Exchange Commission's Staff Accounting Bulletin (SAB) 101 is likely to reduce the frequency and magnitude of such managerial activity.
25. See, for example, Abarbanell and Bernard (1992), Abarbanell and Bushee (1997), and Easterwood and Nutt (1999).
26. Estimated usage growth is computed as of the release date of the first Media Metrix Web Report for the quarter.
27. We extend the trading window to $t = +2$, rather than $t = +1$, because of uncertainty as to whether a firm's earnings announcement is made before or after trading hours.
28. We alternatively use TheStreet.com internet index to compute abnormal returns, and obtain qualitatively similar results to those reported here.

References

Abarbanell, J. and V. Bernard. (1992). "Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior." *Journal of Finance* 47, 1181–1207.

Abarbanell, J. and B. Bushee. (1997). "Fundamental Analysis, Future Earnings, and Stock Prices." *Journal of Accounting Research* 35, 1–24.

Abarbanell, J. and R. Lehavy. (2000). "Differences in Commercial Database Reported Earnings: Implications for Inferences in Research on Analyst Forecast Rationality, Earnings Management, and the Information Content and Value Relevance of Earnings." Working Paper, University of North Carolina and University of California, Berkeley.

Brown, L. (1999). "Managerial Behavior and the Bias in Analysts' Earnings Forecasts." Working Paper, Georgia State University.

Demers, E. and B. Lev. (2001). "A Rude Awakening: Internet Value Drivers in 2000." *Review of Accounting Studies* 6, 331–369.

Easterwood, J. and S. Nutt. (1999). "Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism." *Journal of Finance* 54, 1777–1797.

Hand, J. (2000a). "Profits, Losses, and the Non-Linear Pricing of Internet Stocks." Working Paper, University of North Carolina, Chapel Hill.

Hand, J. (2000b). "The Role of Economic Fundamentals, Web Traffic, and Supply and Demand in the Pricing of U.S. Internet Stocks." Working Paper, University of North Carolina, Chapel Hill.

Ittner, C. and D. Larcker. (1998). "Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction." *Journal of Accounting Research* 36, 1–46.

Kasznik, R. and M. McNichols. (1999). "Does Meeting Expectations Matter? Evidence from Analyst Forecast Revisions and Share Prices." Working Paper, Stanford University.

Matsumoto, D. (1999). "Management's Incentives to Guide Analysts' Forecasts." Working Paper, Harvard University.

Rajgopal, S., S. Kotha and M. Venkatachalam. (2000). "The Relevance of Web Traffic for Internet Stock Prices." Working Paper, University of Washington and Stanford University.

Richardson, S., S. Teoh and P. Wysocki. (1999). "Tracking Analysts' Forecasts over the Annual Earnings Horizon: Are Analysts' Forecasts Optimistic or Pessimistic?," Working Paper, University of Michigan.

Trueman, B., M. H. F. Wong and X. Zhang. (Supplement 2000). "The Eyeballs Have It: Searching for the Value in Internet Stocks." forthcoming *Journal of Accounting Research* 38.