Shrewd, crude or simply deluded? Comovement and the internet stock phenomenon

Ezra W. Zuckerman and Hayagreeva Rao

We show that (i) return comovement among internet stocks during the late 1990s and early 2000 was not exceedingly high; (ii) there was substantial consistency in the manner by which investors distinguished among internet stocks; and (iii) high comovement is most common during periods of price erosion. These results cast doubt on the view that the extraordinary appreciation of internet stocks was due to the crude classification of all such stocks as members of a hot category.

1. Introduction

Managers, investors and policy-makers will be grappling with the aftermath and implications of the rise and fall of the internet stock for many years. Yet perhaps the greatest impact of the internet stock phenomenon may be on scholars due to the sharp slap in the face that the internet stock phenomenon represents to the efficient market hypothesis (EMH). As Ofek and Richardson (2002) point out, the very high volume of trade in internet stocks during their rise in the late 1990s indicates that the seemingly wide gap between the prices they achieved and their fundamental values cannot be easily dismissed as due to illiquidity or some other malfunctioning of the market mechanism. They conclude 'That these prices [were] real is bad news for financial economists, like ourselves, who believe in the rational formation of asset prices' (Ofek and Richardson, 2002: 266).

Two broad explanations have achieved prominence in the emerging literature on how internet stocks reached unjustifiably high prices in the late 1990s and early 2000 (for a review of the insupportable assumptions necessary to justify internet stock prices during this period, see Shiller, 2000; Demers and Lev, 2001; Ofek and Richardson, 2002: 7–12). The first focuses on the nature of speculation in asset markets. In particular, since fundamental values affect capital gains and losses only through changes in market opinion, speculators often focus on indicators of the latter rather than the former (see Keynes, 1960). As a result, market prices may (if only temporarily) be driven by self-fulfilling prophecies (Merton, 1968) or, in the language of contemporary academic finance, they may succumb to 'rational bubbles' (Froot and Obstfeldt, 1991) whereby attention is fixed on prices, which are expected to increase, rather than fundamentals.

A second account for internet stock prices suggests that fundamentals were indeed
responsible for market prices but that the investors interpreting such fundamentals were wildly—even irrationally—optimistic in making their assessments. But why were such 'noise' traders not forced out of the market by more sensible investors, as would be assumed by the EMH? Fundamental limitations on arbitrage (see De Long et al., 1993; Shleifer and Vishny, 1997; cf. Keynes, 1960: 138) may have been responsible. A variant of this interpretation cites limitations on the market’s capacity for short-selling (Miller, 1977; Chen et al., 2001; D’Avolio, 2002; Diether et al., 2002; Jones and Lamont, 2002; Hong and Stein, 2003). Ofek and Richardson (2003) describe a process whereby the significant constraints on the short-selling of internet stocks (due to the relatively small float or shares available for purchase or for borrowing) prevented the opinions of more reasonable investors from being incorporated into prices. Conversely, the fall of internet stocks stemmed from the coincidental expiration, in early 2000, of numerous lock-up periods that had been preventing insiders from selling their shares. These expirations loosened constraints on short-selling, which allowed skeptical investors to give expression to their views, thereby precipitating the subsequent crash in prices (Ofek and Richardson, 2003).

While each of the two candidate accounts of the internet stock phenomenon emphasizes different features of market dynamics, both represent significant departures from the EMH. Whereas the first suggests that fundamental values became decoupled from price because speculators were responsive to one another rather than to indicators of underlying value, the second suggests that investors who focus on fundamentals were removed from the market. According to either story, market prices were not efficient. Indeed, it is striking that financial economists seem more eager to attribute the pricing of internet stocks to inefficiency and even irrationality than are some organization and strategy theorists. Consider two articles published on one of the most discussed ‘anomalies’ of the era: the apparent tendency for investors to bid up the prices of stocks simply for having changed their names to include an internet association such as ‘.com’. Cooper et al. (2001) examine this issue in the Journal of Finance and conclude that:

> Our results are driven by a degree of investor mania—investors seem to be eager to be associated with the internet at all costs . . . this paper adds to a growing body of evidence documenting irrational investor behavior, both at the aggregate and at the individual level. (Cooper et al., 2001: 14)

By contrast, Lee (2001) studies the same phenomenon in a paper published in the Strategic Management Journal and disagrees with the characterization of internet stock investors as irrational. She displays evidence that, over periods of up to a year after the announcement, the market distinguished between companies that changed their names as part of real changes in strategic direction from those that did so for cosmetic purposes only. She concludes that ‘the market is not easily fooled’ (Lee, 2001: 82).

Lee’s analysis is important because it suggests a third approach to the internet stock phenomenon, one that has deeper implications for our understanding of how financial markets function. Rather than being solely the result of a dynamic that encouraged the
disregarding of fundamentals or of constraints that prevented the participation of reasonable investors, perhaps internet stock prices were influenced by an interpretation of fundamentals that was simply wrong. According to this view, investors made distinctions among internet stocks just as they do among other types of stocks in a typical market environment. As with the price of any stock, the marginal investor in internet stocks applied a theory of valuation that maps values of particular business fundamentals onto prices.¹ Several recent papers have demonstrated how various metrics, including some traditional accounting variables and others such as ‘cash burn’, ‘page-views’ and ‘unique visitors’, can be shown to have affected the valuation of internet stock prices during their rise (Trueman et al., 2000; Shivaram et al., 2002). But importantly, the metrics that indicated high valuation before the spring of 2000 were interpreted in opposite fashion afterwards (Demers and Lev, 2002).² Thus, the pricing of internet stocks during the late 1990s may have been driven by a theory (or set of related theories) of valuation that was subsequently debunked. And if so, this analysis suggests that the EMH needs to be modified in a specific but far-reaching way: financial market prices are indeed good estimates of value in light of available information, but they are only as good as the prevailing theories of value that are used to interpret that information (cf. Shiller, 1990; Zuckerman, 1999, 2000). Moreover, such theories are always provisional and it may take years for incorrect theories to be discredited.

In the following, we examine whether internet stocks were priced in a manner that resembled or departed significantly from the pricing of other stocks. We pursue this issue through an analysis of the correlation or ‘comovement’ of stocks within the internet category relative to other categories. The issue of comovement among securities has recently attracted renewed interest (Barberis et al., 2003) because it speaks to the larger question of whether the use of categories by market participants leads them to treat certain pairs of goods or assets as being more alike, and other pairs as more dissimilar, than are their underlying characteristics. If certain assets ‘comove’ more (or less) than they should on the basis of their fundamental or intrinsic value, such ‘segmentation’ would indicate that the market is less sophisticated than imagined by the EMH (Shiller, 1989; Pindyck and Rotemberg, 1993; cf. Lessard, 1974, 1976). Moreover, if crude categorization produces distortionary comovement, it seems most likely to do so in financial markets that are beset by ‘irrational’ optimism or pessimism.

¹We do not mean to imply that all market participants or even the majority are ‘investors’ in Keynes’s (1960) sense—i.e. people who purchase stock for the income (dividend yield) that it generates, as distinct from ‘speculators’, who are sensitive to capital gains and losses (cf. Smith, 1981). But even if the market is dominated by speculators who disregard fundamentals and watch one another or technical indicators, ‘in the end all prices depend on someone’s estimate of future income’ (Williams, 1956: 3). Indeed, prices face at least a soft, lower-bound constraint in that they cannot go too far below the value of its income stream that would accrue to the investor who buys the firm outright and thereby gains access to that cash flow.

²Schwartz and Moon (2000) present another interpretation of internet stocks that sees investors in these stocks as reasonable. They argue that, if considered as real options, internet stock prices could be justified depending on certain assumptions.
For example, various scholars argue that the spread of the Asian financial crisis of 1997 to such distant countries as Brazil, the Czech republic, and South Africa, was driven by investors’ tendency to ignore economic fundamentals and to treat all ‘emerging markets’ securities as members of the same category (Kaminsky et al., 2000).

We analyze comovement among internet stocks during 1996–2000 in an effort to shed light on how investors valued these stocks and, more generally, how financial market prices relate to underlying fundamentals. Accounts that treat internet stock investors as inattentive to or wildly optimistic about fundamentals imply a very high level of comovement among internet stocks. As Cooper et al. (2001) would have it, investors were far from discerning in their selection of internet shares. However, if investors were applying a theory of value that had not yet been discredited, we should see evidence that comovement was not especially high relative to other investment categories.

We present evidence below that, while not definitive, is more consistent with the latter interpretation. In so doing, we describe a methodology for comparing comovement across categories that is novel in that it allows for a common way of identifying the categories in use by market participants such that their dynamics may be compared. This method also allows us to examine two additional features of comovement patterns that have important implications for the question at hand: how a category’s level of comovement in a given period relates to the aggregate return for that period; and how consistent is the structure of comovement among members of a category. Overall, the evidence we present supports our understanding that investment in internet stock prices did not exhibit the crude categorization spurred by speculative excess. Rather, such investment appears to have reflected the dominance of a theory of valuation that was applied consistently, but was fundamentally flawed.

2. Comovement among internet stocks

The heart of our analysis examines the degree of comovement among internet stocks during and after the market reversal, which commenced in the spring of 2000. We use two samples of internet stocks for this analysis. First, we examine internet new issues or initial public offerings (IPOs). Next, we present a common methodology for sampling internet stocks as well as all other major industry categories, which allows us to compare the level of comovement among internet stocks with that displayed in other industries.

2.1 Comovement among internet IPOs

Our sample of internet IPOs consists of the 451 (non-subsidiary) companies in the

---

3The timing of the ‘crash’ is hard to pin down. The intraday NASDAQ peak of 5132 was reached on March 10, 2000. However, the decline in the index for the month of March was not substantial (from 4697 to 4573, or –2.6%). On a monthly basis, the decline did not begin in earnest until April (close of 3861, or a –16% monthly return).
SDC database that: (i) went public during the 5 year period beginning in January 1995 and ending in December 1999; (ii) highlighted the internet in the description of their business models; and (iii) were included in the CRSP database, from which we obtain post-IPO data on stock price and returns. Using a similar sample selection process over an observation window that was slightly smaller than ours (January 1996 through March 2000), Schultz and Zaman (2001) found a similar number (420) of stocks. Figure 1 shows the number internet IPOs in each month over the sample period.

Note that a sample of internet IPOs is not necessarily a sample of 'internet stocks'. First, there are stocks that went public before 1995 and/or altered their business models such that they became heavily oriented towards the internet. America Online, which began as a proprietary service and then evolved to include access and services related to the internet, is perhaps the best example of such a stock. However, since it is highly unlikely that our sample of internet IPOs includes firms that reoriented their strategies away from the internet during this period, we can be confident that it does not include the stocks of firms with business models that were not internet related. A second concern is that the internet designation is too broad and that there may be subcategories (e.g. e-commerce stocks, business-to-business, internet infrastructure) that were more coherent. The second sampling strategy we describe below addresses this issue.

In Figure 2, we plot the two trend lines that describe stock price patterns in our sample from the third-quarter of 1995 (by which time, there were ten stocks in our

Figure 1 Month of IPOs in sample.

---

38Thirty-eight stocks were eliminated on the last selection. Most of these appear to be American Depository Receipts (ADRs) of foreign companies.
sample) through the end of 2000. One trend line describes the mean monthly return for the sample stocks.\textsuperscript{5} We exclude from the monthly samples any stock that went public in that month, so as to remove any patterns that are unique to the going-public process. As we can see, the return on the sample stocks was quite high through February 2000 (mean monthly return of 9.1%) and quite negative over the following nine months (mean of \(-14.4\%\)).\textsuperscript{6}

The second trend line in Figure 2 describes the mean correlation in the daily return for sample stocks during a given month. For each month, we compute the correlation between the daily return vectors for each pair of stocks in the sample. We then compute the mean correlation taken over all pairs in the monthly sample. As in prior work on stock comovement (e.g. Shiller, 1989; Pindyck and Rotemberg, 1993; Barberis \textit{et al.}, 2002), we regard a high correlation between a pair of stocks as indicating that stock market participants treat the two firms as fundamentally alike. At the extreme of a correlation of 1, information that arrives on the market relevant to the valuation of either stock will induce movement by both stocks in the same direction and to the same degree. By contrast, if a pair of stocks display a negative correlation, this would indicate

\textsuperscript{5}The trend line for mean return weighted by outstanding shares looks virtually the same. The correlation between the two vectors is 0.97. We display the unweighted mean to simplify the presentation.

\textsuperscript{6}Each of these numbers is considerably higher on a compounded basis; however, computing cumulative returns is not straightforward since our sample changes in its composition from month to month.
that market participants interpret information regarding one as having the opposite implications for the other. No correlation between the two stocks would suggest the application of orthogonal theories of value to the two securities. In taking the mean correlation among a group of stocks, we may assess the extent to which market participants regard members of the group as fundamentally alike one another.\footnote{Other metrics for assessing the similarity of vectors could be used. Possible problems with the correlation coefficient for current purposes include: (i) it eliminates differences in the mean values, which could be of interest; (ii) it focuses on deviations from mean values but ignores deviation from the origin, which might be important if one regards positive and negative returns to be qualitatively different; and (iii) in conditioning daily variation on the mean for a month, it effectively characterizes the present in terms of the future, which seems awkward. The first issue does not concern us in the present case because we are interested not in differences in level but in pattern or shape. Moreover, use of a metric such as Euclidean distance would be problematic because differences in level will overwhelm differences in shape, especially in aggregate comparisons. In particular, since the mean return of internet stocks was so high during much of this period, the mean distance between vectors in this category will be high as well, resulting in very low comovement. The second and third issues are somewhat more troubling and so we performed a robustness check by re-running all our models with an alternative metric: for each pair of return vectors, we computed the proportion of trading days in a month in which both stocks moved in the same direction—up, down or no change. This metric avoids both problems (ii) and (iii), though it removes differences in the size of change: a 1% positive return is as much of a match with a 0.1% positive return as it is with another 1% positive return. Results using this ‘directional matching’ metric were virtually the same as those using the correlation coefficient.}

As reviewed above, each of the two dominant accounts of the internet stock phenomenon would suggest that investors treated internet stocks in a rather indiscriminate manner, which is consistent with a very high level of comovement in this category. Furthermore, this literature suggests that such a tendency to crudely lump all internet stocks together should be particularly evident during periods of rapid appreciation when investors presumably had insatiable demand for ‘all things internet’. By contrast, investors should have been more discerning during the subsequent ‘crash’ period.

This is not what we find in our sample of internet IPOs. As we see in Figure 2, the mean monthly correlation among sample stocks ranged from 0.01 (July 1997) to 0.37 (April 2000), with a mean of 0.010. It seems fairly clear that at no point during either the rise or the fall of internet stocks did investors treat these stocks as being substantially the same. Moreover, comovement seems more characteristic of price drops than of rises. Accordingly, while the three most notable positive spikes in mean monthly return (November 1998: 57.3%; November 1999: 43.8%; and April 1996: 36.9%) were not associated with high comovement (mean correlations of, respectively, 0.07, 0.05 and 0.07), the three months with the largest downward movement (November 2000, –34.4%; April 2000, –28.0%; August 1998, –28.0%) were among the highest in their level of comovement (0.17, 0.37 and 0.25, respectively). Indeed, the mean monthly correlation averaged 0.09 through February 2000 versus 0.13 thereafter. The patterns presented in Figure 3, which plots the relationship between mean monthly return and mean monthly return correlation, reinforces the conclusion that a high level of

Shrewd, crude or simply deluded? Comovement and the internet stock phenomenon 177
comovement is more characteristic of drops in market value rather than it is of price rises and suggests a nonlinear relationship. Over most of the range of returns, there is little relationship between the two variables. However, returns for the monthly sample of below −20% and particularly below −30% are characterized by particularly high levels of comovement.

Thus, the results we have displayed thus far cast doubt on the assumption that internet stock investors were crudely lumping all internet stocks together, particularly during the rapid rise of such stocks. Indeed, the highest levels of comovement are in evidence during significant drops in the value of internet stocks, particularly after February 2000, when more reasonable investors would seemingly have had the predominant influence on valuations (Ofek and Richardson, 2003). But this analysis raises several questions. First, how does comovement among internet stocks compare to the level of comovement among members of other industry categories? After all, it may be that the observed level of comovement is relatively high even if it looks rather low in absolute terms. Moreover, perhaps the level of comovement among internet stocks is relatively high compared to the underlying coherence in the economic fundamentals of the constituent stocks. To the extent that the internet is simply a technological platform for a wide variety of business models, it may be argued that there should be little comovement among these stocks in some fundamental sense such that the observed level of comovement is in fact quite high. Finally, we would like to investigate the possibility that the seemingly low level of overall comovement masks substantial comovement among subsets of the category. Our second sampling strategy addresses each of these issues.

Figure 3  Mean monthly return and return correlation: internet IPOs, October 1995 through December 2000. \( y = 0.0794 - 0.118x + 0.872x^2 - 1.2594x^3; R^2 = 0.3549. \)
3. A method for comparative comovement analysis

While our first sampling strategy relies on self-categorization by firms at the time of IPO, our second approach selects stocks based on how key market participants classify such stocks during our investigation window. The classifications we use emerge implicitly from the relationship between large institutional investors and the universe of equities, as mediated by Wall Street securities analysts. As Zuckerman (1997, 1999, 2000, 2004) has shown, securities analysts’ specialization by industry provides an opportunity for investigating the implications of industry-based stock market categories for market dynamics. Since analysts are typically responsible for one and sometimes two related industries, an analyst’s coverage of a particular stock can be taken as an implicit classification of that stock as being a member of the analyst’s industry. Thus, for example, if one wants to see which stocks are internet stocks, one looks to see which stocks are covered by internet analysts.

However, while this procedure is useful when firms define their industries in terms of official industry classifications (SIC or NAICS codes; e.g. Zuckerman, 2000), it is less helpful in cases such as the internet and other (new) industries that are not associated with a clear official code. How do we decide who is an internet stock analyst? And how can we make such determinations in a way that can be replicated across the full set of industry categories? We address these issues by measuring implicit classifications of analysts via their rankings by institutional investors.

As do Phillips and Zuckerman (2001: 406–407), we use the raw survey data that Institutional Investor magazine (II) collects in the spring of each year and publishes in summary fashion in its October issue. The sampling universe consists of the largest ‘buy side’ money managers—for example, banks, insurance companies, pension funds and mutual funds, a list that includes the members of the 300 largest institutional investors, as well as additional significant institutions obtained from industry sources. Questionnaires are mailed out in April for return in June. Respondent institutions are encouraged to give responsibility for filling out the questionnaire regarding a particular industry (category) to those individuals (typically, buy side analysts) who best represent the institution’s views regarding that category. We utilize poll data from each of the five years from 1996 to 2000. The response rate for each of these years was quite high. In 1996, for example, 1337 individuals responded for 300 firms. Among the respondents were 68% of the II 300 and more than 90% of the Institutional Investor 100.

The main survey question that is used to determine the analysts’ relative ranking in a given industry is the following: ‘We are putting together our ‘team’ of outstanding brokerage analysts for (year) and would like your help. Please rank in order your

---

8Several categories are not industry-based. For instance, in 2000 there were categories for Accounting and Tax Policy, Quantitative Research, Small Companies, and Washington Research.

9II does not reveal the exact response rate for categories that exceed 90%, lest the identity of respondents be deducible.
selections for the best analysts . . . during the past twelve months. Respondents are then given the opportunity to list four analysts for a given category, in rows ranked one through four. For the present analysis, we are unconcerned with an analyst’s overall ranking. Rather, we exploit the fact that a ranking by a survey respondent of an analyst represents an implicit classification of that analyst as a specialist in that industry. The more numerous are such implicit classifications, the greater the association between the analyst and the industry in the minds of key market participants. In particular, we treat an analyst as covering a particular industry when the analyst is nominated by at least five respondents as one of the ‘best analysts’ in that area. And once having established which analysts specialize in a given category, we are then in a position to measure which stocks are classified in that category based on coverage by specialists. That is, to the extent that a given stock is covered by analysts who are implicitly classified as industry specialists by II respondents (and it is not covered by analysts who specialize in other categories), we may presume that it is regarded by market participants as a member of that category.

We clarify the procedure by considering two industries: ‘Internet & New Media’ (INM) and Airlines. In 1999, fifty-four different analysts received at least one nomination as a top INM analyst and twenty-five were nominated as a top Airline analyst. As with all such rankings (see Phillips and Zuckerman 2001), the allocation of nominations was highly skewed. Thus, thirty-two of the INM and fifteen of the Airline analysts cited received only one citation, while the top analysts received 103 and eighty nominations, respectively. In all, seventeen INM and ten airlines analysts received at least five nominations. As is typical, the analysts who received at least five nominations each represented a different, high-status, ‘bulge bracket’ firm. The analysts with fewer citations were generally of three types: (i) junior or assistant analysts at bulge bracket firms; (ii) analysts who represented lower status firms, which often have a regional or specialized focus; and (iii) high ranking analysts in other industries whose range of coverage touches on the focal industry in a peripheral way. We disregard the analysts who received fewer than five nominations because we do not believe that these three types of analysts should be regarded as specialists in a given category. In the following, we refer to those analysts who received at least five nominations as ‘II analysts’ and we refer to the II analysts for a given industry as ‘II specialists in that category.

The coverage provided by II analysts provides a basis for classifying stocks in various industries. We code an analyst as covering a stock during a given period when the analyst publishes an earnings estimate in the IBES database during that period. IBES data cover the vast majority of earnings estimates and thus are an accurate representation of the pattern of coverage. In the second quarter of 1999 (the quarter

---

180 E. W. Zuckerman and H. Rao

---

10We tested this assumption by searching for non-II analysts in the IBES database who were structurally equivalent to (i.e. high degree of overlap in coverage with) the II specialists in a particular industry. Typically, there are very few such additional specialists in a given industry and inclusion of such analysts does not change the results. We decided to use only the II analysts for our measures to simply the procedure.
during which the 1999 *II* survey was conducted), fifteen of the seventeen *II* INM analysts and eight of the ten of the *II* Airlines analysts are represented in the IBES data.\(^{11}\)

In Tables 1 and 2, we list the stocks that were covered by at least two INM or Airlines specialists and we note the total number of *II* analysts (i.e. including those who specialize in other industries) who cover these stocks. As is evident, these lists seem to capture their respective categories quite well.

Our procedure for classifying stocks emerges implicitly from the interaction between large investors, analysts, and stocks and thereby should be an accurate representation of how market participants categorize stocks. In addition, the procedure leads to a recognition that stocks *vary* in the extent to which they may regarded as representative of a category. Such variance is largely a reflection of the stocks’ relative sizes and their strategies. Thus, America Online was clearly more of an INM stock than was Microsoft, which appears in Table 1c largely because some *II* INM analysts are also *II* Software analysts, all of whom covered Microsoft. In general, it will make sense to regard a stock as a member of a category to the extent that: (i) it receives coverage from many *II* specialists in that category; and (ii) such analysts represent a large fraction of the total coverage it received from *II* analysts.

Our method also helps us address the three issues raised above with respect to our sample of IPOs. First, we now have a common procedure for generating samples of stocks across all industry categories, which allows us to compare the level of comovement among constituent stocks. Specifically, we may analyze the question as to whether comovement among INM stocks was high or low. Second, since our procedure directly measures how investors and analysts categorize stocks, we may assess the degree of heterogeneity that pertains to a given sample of presumably homogeneous stocks, such as our sample of internet IPOs. In Table 3, we present the distribution of industry specialties of the *II* analysts that covered the internet IPOs in each of the second quarters from 1996 through 2000. Each observation in this table represents coverage of a member of our IPO sample by an *II* analyst that specialized in the industry category.\(^{12}\)

Note first that about half of the internet IPOs do not appear on this graph because they were not covered by any *II* analysts during these periods. In particular, 22% (6 of 27) of the sample did not receive coverage by an *II* analyst in the second quarter of 1996, while 53% (42 of 79), 50% (61 of 122), 36% (69 of 192), and 46% (202 of 436) failed to receive such coverage in the second quarters of 1997, 1998, 1999 and 2000, respectively. Thus, our procedure does not indicate how such (smaller) stocks were categorized by

---

\(^{11}\)These rates of exclusion of *II* analysts from the IBES database are typical. The mean rate of exclusion for the quarterly samples was 15.5%, with a high of 20% (1996, third quarter) and a low of 12.6% (1999, second and third quarters).

\(^{12}\)We code instances where the same analyst is nominated as an *II* specialist in more than one industry as representing more than one ‘speciality-stock’ pair. We present only those specialities that represented at four observations in at least one of these time periods. We also collapse several categories (e.g. PC and Server Software) to facilitate presentation.
market participants because these stocks did not receive sufficient attention from II analysts.

The table reveals substantial heterogeneity in the internet IPO sample in terms of their categorization via coverage by II analysts. In particular, while many of the IPOs were covered by INM specialists, roughly comparable properties in a given period were covered by Software specialists. That is, our IPO sample includes several companies that described their business strategies at the time of IPO as involving the internet but were

Table 1 Stocks with coverage by two or more II Internet and New Media (INM) specialists, second quarter of 1999

<table>
<thead>
<tr>
<th>Name of stock</th>
<th>Number of II internet analysts covering</th>
<th>Total number of II analysts covering</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Stocks with coverage by 4+ II INM specialists</td>
<td></td>
<td></td>
</tr>
<tr>
<td>America Online</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Yahoo!</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Lycos</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Amazon.com</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Excite</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>At Home</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Cnet</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Infoseek</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Doubleclick</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Ebay</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Priceline</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>SportsLine USA</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>(b) Stocks with coverage by 3 II INM specialists</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autobytel.com, Beyond.com, Broadcast.com, Digital River, IXL Enterprises, iVillage, Starmedia Network, The Street.com, Ticketmaster Online/Citysearch, Xoom.com, Ziff Davis Inc.</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Checkpoint, Security Dynamics, Sterling Commerce</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Electronic Arts</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(c) Stocks with coverage by 2 II INM specialists</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At Plan, Autoweb.com, Careerbuilder, Intuit, Intervu, Juno Online Services, Media Metrix, Network Solutions, Open Market, Onsale, Preview Travel, Secure Computing CMGI, Harbinger Inc., Real Networks, TMP Worldwide</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>E*Trade Group</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Pixar</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Microsoft</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>
categorized by market participants as software companies. Notable examples (and the number of II Software specialists covering them in the second quarter of 2000) are: Ariba Inc. (sixteen Software specialists), Legato Systems (fourteen) and Siebel Systems (ten).

We also see the emergence of several new categories in the later years, with Internet Infrastructure the most notable. In sum, given the observed level of heterogeneity, it should be less surprising that we find relatively little comovement among stocks in our IPO sample. We thus shift our analytical tack to analyze subsamples of stocks based on how they are implicitly classified by market participants. It seems reasonable to expect higher intra-category comovement within such subsamples than was found in our IPO sample.

### 3.1 Categorical coherence

Before describing the results from this analysis of comparative comovement, we highlight an additional advantage of our procedure. As noted above, any assessment of the degree of comovement among a set of stocks should be made in relation to the level expected given relatedness in the economic fundamentals (firm characteristics and

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Stocks with coverage by two or more II Airline specialists, second quarter of 1999</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name of stock</strong></td>
<td><strong>Number of II Airline specialists</strong></td>
</tr>
</tbody>
</table>

#### (a) Stocks with coverage by 4+ II Airline specialists
- American Airlines: 7
- United Airlines: 7
- Continental Airlines: 7
- Northwest Airlines: 7
- Southwest Airlines: 7
- Delta Airlines: 6
- US Airways Group: 6
- Alaska Airlines: 5
- America West Airlines: 5
- Skywest Airlines: 4

#### (b) Stocks with coverage by 3 II Airlines specialists
- Atlantic Coast Airlines, Midwest Express, TWA: 3
- Comair Holdings: 4

#### (c) Stocks with coverage by 2 II Airlines specialists
- Mesa Airlines, Airtran Holdings, Lan Chile S.A.: 2
- Amtran, Mesa Air Group: 3
- Atlas Air: 5
environmental factors) that pertain to those stocks. 13 One possible approach to this problem would involve using similarity coefficients based on accounting data. Yet this is problematic because different accounting data are used in various industries, thus making cross-industry comparisons difficult. Moreover, it is not clear how to treat stocks for which the most relevant variables (e.g. dividends) are not available.

### Table 3a. Fixed-effects regression of comovement on return: January 1996 through September 2000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of stocks/100</td>
<td>-0.193 (0.032)</td>
<td>-0.193 (0.031)</td>
<td>-0.210 (0.031)</td>
<td>-0.208 (0.030)</td>
</tr>
<tr>
<td>Mean Return</td>
<td>-0.327 (0.047)</td>
<td>-0.305 (0.047)</td>
<td>-0.349 (0.040)</td>
<td>-0.349 (0.040)</td>
</tr>
<tr>
<td>Mean Return²</td>
<td>0.947 (0.074)</td>
<td>1.447 (0.108)</td>
<td>1.447 (0.108)</td>
<td>1.447 (0.108)</td>
</tr>
<tr>
<td>Mean Return³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.236 (0.007)</td>
<td>0.236 (0.007)</td>
<td>0.237 (0.007)</td>
<td>0.232 (0.007)</td>
</tr>
<tr>
<td>R²</td>
<td>0.442</td>
<td>0.460</td>
<td>0.494</td>
<td>0.497</td>
</tr>
<tr>
<td>n-observations (industries)</td>
<td>3813 (101)</td>
<td>3813 (101)</td>
<td>3813 (101)</td>
<td>3813 (101)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. All coefficients are significantly different from 0 at \( P < 0.001 \).

*Each model includes a dummy variable for each industry, excluding one industry as a reference category.*

### Table 3b. Fixed-effects regression of adjusted comovement on return: January 1996 through September 2000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean return</td>
<td>-0.334 (0.126)</td>
<td>-0.511 (0.135)</td>
<td>-0.518 (0.135)</td>
</tr>
<tr>
<td>Mean return²</td>
<td>2.124 (0.575)</td>
<td>2.058 (0.842)</td>
<td></td>
</tr>
<tr>
<td>Mean return³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.008 (0.011)</td>
<td>-0.008 (0.012)</td>
<td>-0.160 (0.013)</td>
</tr>
<tr>
<td>R²</td>
<td>0.502</td>
<td>0.504</td>
<td>0.504</td>
</tr>
<tr>
<td>n-observations (industries)</td>
<td>3801 (101)</td>
<td>3801 (101)</td>
<td>3801 (101)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. All coefficients are significantly different from 0 at \( P < 0.001 \).

*Each model includes a dummy variable for each industry, excluding one industry as a reference category.*

*bNot significant. All other coefficients are significantly different from 0 at \( P < 0.001 \).*
We therefore adopt a useful indirect approach, which involves examining the extent to which a set of stocks is considered by securities analysts to be members of the same category. The following assumption underlies this method: a pair of stocks is more likely to be grouped together by analysts as members of the same category to the extent that the stocks are subject to the same set of economic fundamentals. This assumption is based on our general understanding of classificatory structures as devices organizing the interpretation and valuation of a field of objects according to a shared theory of value. For example, the way a society sorts animals according to their legitimate uses—e.g. to serve as companions, to be eaten, to be worshipped—is governed by the cosmology that dominates the culture (Douglas, 1966). Similarly, the categories used in a labor market reflect the theories of skill that imply distinctions between different types of work (Zuckerman et al., 2003). So too, the stock market’s division into industry-based sectors reflects the basic theory that a firm should be valued according to its expected income stream (see Burk, 1988) coupled with the premise that different industries involve distinct forces that shape such earnings prospects (Zuckerman, 1999, 2000, 2004). This does not mean that the classifications are ‘correct’ but that they are generally found useful as devices for applying the general theory of value.

Of course, that industry classifications guide market participants in distinguishing stocks from one another on the basis of their fundamental drivers, does not mean that all categories are equally useful in this regard. In particular, categories themselves vary in their level of ‘coherence,’ as reflected in the tendency for the analysts who specialize in that category to cover the same sets of stocks (cf. Zuckerman, 2004). At the extreme of high coherence, the analysts who specialize in an industry all follow the same stocks; at the other extreme, there is no overlap in the stocks they cover. Two factors are important in determining such variation in structural coherence. First, while a common label may be applied by analysts to a set of stocks, such stocks may be more or less heterogeneous in their fundamental drivers of value when compared with other categories. A second factor is the relatedness among categories in the classificatory system. If a category is truly distinct in terms of its economic fundamentals, it should be covered by specialists who do not provide coverage on other stocks since the unique aspects of the industry would require highly specialized analysts (cf. Zuckerman et al.,

numbers. Here we intend fundamentals to mean the underlying economic factors that determine its capacity for returning cash to its owners (i.e. in the form of dividends). While there is agreement on the ‘dividend discount model’ (which says that an asset is more valuable the greater and the more proximate in time are the dividends that it returns to its owners) as the theoretical basis for a firm’s intrinsic worth there has never been agreement on what accounting or other indicators might predict such income patterns. See Zuckerman (2004) for discussion on these and related points.

14There are limits to how incoherent a category may be, at least in our data. Analysts are ultimately beholden to (institutional investors). If investors do not attribute a common category to a set of analysts (who would presumably be using different names to describe their specialty), they will not be able to answer the survey. Indeed, Institutional Investor magazine engages in a lengthy process during the months subsequent to the fielding of the survey to ascertain the prevailing categories in investment banks and in the eyes of institutional investor (D. Murrell and D. Blaney, personal communication).
2003). However, if two categories have similar economic fundamentals, they should be followed by similar analysts.\textsuperscript{15}

To return to our case, we might reasonably suspect that INM was a fairly incoherent category. After all, the internet hardly represents an industry in the traditional sense of occupying a relatively clear position in an economic value chain and having relatively identifiable sets of buyers, suppliers, entrants, substitutes and complements. Rather, the internet seems more properly construed as a technological platform for a wide variety of business models. Moreover, the economic fundamentals of internet firms clearly overlapped with those of other categories in information technology and telecommunications, which should also make it a less coherent category. And if this is so, it would again be unsurprising that there would be relatively low levels of comovement among members of the category. The question then becomes whether the level of comovement is high relative to the level expected given the coherence of the category.

How coherent was the INM category during the late 1990s and early 2000? The data summarized in Figure 4 provide an answer. We measure coherence following the method developed by Zuckerman (2004) at the level of the individual stock. First, we divide our sample into three-month periods corresponding to calendar quarters. Second, we define an analyst as covering an industry $i$ during quarter $q$ if the analyst received at least five nominations by $II$ respondents in the poll that was completed: (i) in June of the same year for the first, second and third quarters; or (ii) in June of the next year for the fourth quarter.\textsuperscript{16} As mentioned above, we define an analyst as covering a stock during the quarter if she published at least one earnings report or made one stock recommendation on that stock during the quarter.

Using these sampling criteria, we first define a matrix $M_{iq}$ for industry $i$ during quarter $q$. The dimensions of the matrix are $a \times f$, where $a$ indexes $II$ specialists in industry $i$ during $q$ and $f$ indexes the firms that received coverage by at least one such

\textsuperscript{15}Indeed, at the extreme, the same set of analysts covers each category. The categories ‘Autos’ and ‘Auto Parts’ are a good example. In each year of our sample, overlap in the specialists in each of these categories is very high; at least half of the $II$ specialists in one category are also recognized in the other industry. We treat such instances of high overlap as if they are a single category. In particular, we collapsed two (or more) categories when, in most years, at least half the analysts who specialize in one category are also specialists in the other category. This procedure was also applied to internet stocks in 2000. For that year’s survey, $II$ split the INM category (named the ‘Internet’ category in 1996 and 1997) into two subcategories: ‘E-commerce’ and ‘New Media’. However, we collapsed these two categories because they displayed high overlap in their $II$ analysts. Of the fourteen analysts who received at least five nominations as top E-commerce analysts, nine were also nominated as top New Media analysts. Of the eighteen top New Media analysts, nine were ‘E-commerce’ analysts.

\textsuperscript{16}Since the poll asks respondents to rank analysts on the basis of the prior 12 months, it seems inappropriate to apply the results from June to the subsequent quarter. However, preliminary analysis of the data suggests that the poll appears to reflect the more recent quarters than the more distant ones. For example, analysts who left the profession in the summer tend not to be mentioned in the following spring’s poll. Supporting this assessment, the trend lines for coherence show apparent drops in the third quarter if the alternative measurement scheme is used (this is as true for internet stocks as for other categories).
Figure 4 Quarterly distributions of categorical coherence. Mean number of stocks followed by each pair of II specialists in the industry. Outliers are black dots. Asterisks (linked) are coherence for Internet & New Media.
specialist during \( q \). Based on the \( M_{ij} \) matrix, we then create a matrix \( O_{ij} \) of dimensions \( a \times a \), which measures the degree of overlap in coverage displayed by the specialists in \( i \) during \( q \). The cells of this matrix, \( o_{ij} \), can be measured using any metric of binary proximity. We have experimented with several such metrics, all of which produce similar results. For present purposes, we use the simplest, which is the number of firms that both analyst \( i \) and analyst \( j \) follow (which may be generated by multiplying \( M_{ij} \) by its transpose; see Breiger, 1974). To generate a summary measure of coherence, we take the mean score across all pairs of analysts:

\[
c_{ij} = \frac{1}{A} \sum_{a \neq i} o_{ia} (A \cdot [A - 1]/2)
\]

In Figure 4, we observe how categorical coherence varies across industries and over time. The overall distribution is quite consistent throughout our sample. In every quarter, the typical pair of \( II \) specialists covered roughly four stocks in common with about half the industries in the range from about three to six stocks in common. As is readily apparent, INM was a rather incoherent category in the first years of its existence, but became steadily more coherent such that, by early 1998, it was near the median level. But there is then a noticeable drop in INM’s level of coherence in the final four quarters. This downturn may result from increasing de-differentiation in the category in the later period.\(^{17}\) These patterns suggest that, as one might expect given the strategic heterogeneity encompassed by the internet, INM was a relatively incoherent category during most of this period. More important, we now have a period-specific fundamental baseline against which we may compare the degree of comovement we find with the coherence of the investment category. To the extent that categorical coherence predicts high comovement, we should not be surprised if members of a relatively incoherent category, such as INM, exhibit low comovement.

4. Analysis of comparative comovement

We now compare the raw level of comovement among INM stocks with that observed among stocks in other categories. In Figure 5, we depict how the mean pairwise return correlation (hereafter, ‘comovement’) is distributed across all industries in each of the months from January 1996 through September 2000. We also highlight comovement among INM stocks during these months. In this and subsequent analyses, we include a stock in a category for a given quarter when: (i) it received coverage by at least two \( II \) specialists in that category during the quarter; and (ii) this represented the maximum coverage it received from specialists in any category. For example, all stocks listed in Table 1 are included as members of the INM category for each of the three months in

\(^{17}\)It may also reflect the fact that an unusually high proportion of \( II \) specialists in INM left the profession in 1999 to become venture capitalists and money managers.
the second quarter of 1999, except for Microsoft and Pixar, which are categorized in
other categories (Software and Entertainment, respectively) because they received more
coverage from II analysts in the other categories than they received from II INM
analysts. We exclude from our analyses those stocks that received equal coverage from
analysts in more than one industry.

As we can see, INM stocks did exhibit elevated comovement in several of these
months. In particular, INM was in the fourth quartile of the comovement distribution
in 18 of these 57 months. The most notable periods of high comovement were January
and February 1996, during which INM displayed the second highest (January) and
highest (February) comovement; and August 1998 through June 1999, during which
the average rank of INM was seventh (of seventy industries with at least 9 months of
data).

However, while the level of comovement observed for INM was high during this
11 month period, it was not especially high overall. The INM category’s rank from
January 1996 through December 1999 (2000) was only twentieth (twentieth) of
seventy-two (sixty-one) industries with at least 2 (3) years of data. This ranking is
hardly what is implied by accounts that depict internet stock investors as failing to
discriminate among members of the category. Moreover, even between August 1998
and June 1999, there were six industry categories that displayed greater comovement
[Oil Services & Equipment, Pharmaceuticals/Major, Airlines, Brokers & Asset
Managers, Banks, Oil (Major) and Airlines], none of which was new or high-tech
related such that they could be said to be driven by crude categorization. Rather, the
most straightforward explanation for high comovement within such categories is that
the determinants of their fundamental value are relatively homogenous.

Yet it might still be the case that INM stocks comoved at a high level when one
controls for the coherence of the category. As we observed in Figure 4, INM was a
relatively incoherent category through much of this period. Moreover, there is a clear
relationship between categorical coherence and comovement. In Figure 6, we plot a
category’s mean comovement against its mean coherence for those sixty-six categories
that had at least 3 years of data. The correlation between these two variables was 0.45,
which reinforces our treatment of coherence as a baseline against which the category’s
comovement may be assessed. The plot reveals that INM did indeed exhibit a level of
comovement that was high relative to what would be expected on the basis of the
category’s coherence. However, it also suggests that if comovement among INM stocks
was excessive, it was also excessive in such categories as Banks, Tobacco, Nonferrous
Metals, Semiconductors, Brokers and Oil Services. The most straightforward
interpretation of the high residuals for these categories is that our measure of
categorical coherence underestimates the extent to which the economic fundamentals
of constituent stocks are interrelated—not that investors were engaging in crude
categorization. And if so, this would mean that, if comovement among internet stocks
looks high, it may be because investors paid attention to fundamentals (but perhaps
used an incorrect theory in doing so) rather than ignored them.
Figure 5  Monthly distributions of mean comovement. Mean pairwise correlation of daily return taken over all pairs of stocks in the industry. Outliers are black dots. Asterisks (linked) are mean comovement for Internet & New Media.
Figure 6 Mean comovement by mean categorical coherence. Values taken over all months for the seventy-one industries with 2+ years of data. Regression equation ($R^2 = 0.198$): mean comovement $= 0.091 + 0.025 \times$ mean categorical coherence.
In Figure 7, we break down these patterns by month. For each month, we regress a category’s comovement on its categorical coherence for the quarter as well as the number of firms in the category that month (comovement should be higher in a category with fewer stocks) and calculate the studentized residuals from this equation. We display the distributions of these residuals (hereafter, ‘adjusted comovement’) in Figure 7, with INM’s adjusted comovement highlighted. As compared with Figure 5, INM’s relatively high comovement is clearer, especially for the 24 month period beginning in September 1997 and ending in August 1999. During these 2 years, INM was either in the third or fourth quartile of the distribution in every month and more commonly the latter (21 of 24 months). Thus, it seems clear that INM did exhibit a level of comovement during this period that was quite high when compared with other categories and when judged against the underlying coherence of the category.

Yet the patterns do not readily accord with the prevailing view that investors crudely grouped all internet stocks together as befit their mood of ‘irrational exuberance’ (Shiller 2000). As with the patterns for unadjusted comovement, INM does not exhibit the highest adjusted comovement of any category, even for the 2 year period between September, 1997 and August, 1999. INM was eighth (of sixty-two categories with the full 2 years of data) in mean adjusted comovement (mean of 0.94 standard deviation units above the monthly regression lines) during this period, after the following industries: Oil Services & Equipment, Pharmaceuticals (Major), Brokers & Asset Managers, Tobacco, Oil (Major) and Nonferrous Metals. Again, the most straightforward interpretation of these industries’ high levels of adjusted comovement is that categorical coherence underestimates the extent to which the economic fundamentals of members stocks are related. And if we thereby treat these categories as exhibiting a moderate level of comovement, the same interpretation would seem to apply to INM. Moreover, INM did not exhibit as high a level of adjusted comovement during the rest of our window of observation. It was sixteenth (out of seventy-two categories that had at least 2 years of data) from January 1996 through December 1999, and eighteenth (out of sixty-one categories with at least 3 years of data) from January 1996 through September 2000.

4.1 Comovement and return

Thus, while we have found that stocks in the INM category exhibited a high degree of comovement during portions of our sample, it is questionable whether the comovement achieved the excessive levels suggested by prominent accounts of the internet stock phenomenon. Still, the evidence we have presented is not conclusive. It remains possible to argue that the level of comovement observed, particularly from mid-1997 through mid-1999, was unjustifiably high and that it reflected a speculative environment in which investors were not as discriminating as they should have been given the underlying fundamentals. We have given some reason to doubt this conclusion, particularly by pointing out that the industries that exhibit the highest
Figure 7  Monthly distributions of adjusted comovement. Studentized residual from monthly regression of mean pairwise return correlation of number of observations and categorical coherence. Outliers are black dots. Asterisks (linked) are mean coherence for Internet & New Media.
adjusted comovement are those that are probably the most coherent on a fundamental basis.

There are two additional reasons to doubt the assertion that internet investors were 'exuberantly indiscriminate'. First, as we observed in Figure 3 for our IPO sample, high comovement appears to be more common during periods when stocks in a given category are losing value. In Table 4, we present results from fixed-effects regression analyses of comovement on mean return and the number of stocks for each month, from January 1996 through September 2000. Each model includes a series of dummy variables to capture baseline industry effects. The variance explained is thus modeled in relation to a global constant and a constant for each industry. In Table 4a, we present results for comovement, with those for adjusted comovement in Table 4b. The patterns in the first table are similar to what we observed in Figure 3. That is, comovement appears to be highest during periods of very low returns, to be unrelated to returns over the intermediate range of the latter, and then to be lowest at the highest range of returns. In Figure 8, we graph the relationship between mean monthly return and adjusted comovement, as indicated by model 2 in Table 4b. As the semi-parabolic shape indicates, high adjusted comovement may characterize periods of high, positive returns as well as low. Yet the main pattern that emerges is the strongly negative association between return and comovement.

In Figures 9–11, we illustrate these results with the three categories that are most relevant to any analysis of the internet stock phenomenon: INM, Internet Infrastructure and Telecom Equipment (Wireline). Several patterns are evident in Figure 9. First, we see that INM stocks follow a similar trajectory as the IPO sample in terms of monthly appreciation, though the descent begins earlier for the former. While 26 of the 45 months through December 1999 exhibited a positive return (mean of mean monthly returns = 0.064), 9 of the 12 months of 2000 showed negative returns (mean of mean monthly returns = –0.112). Second, comovement among INM stocks is considerably higher than that observed for the more heterogeneous IPO sample. The mean comovement across all months is 0.21, which is more than twice the level observed among the IPOs.

One apparent difference between INM stocks and the IPO sample is that, among the former, comovement is typically higher before the descent (mean comovement through December 1999 = 0.217) than beforehand (mean comovement during 2000 = 0.171). However, this aggregate difference masks the tendency for high comovement to occur during months of significant price declines throughout our observation window. In particular, we see that the months that displayed the highest comovement and adjusted comovement were almost always months with rather low returns. Notable examples include January and February 1996, October and November 1997, August 1998 and February 1999.

The patterns in Figures 10 and 11 are even clearer. As might be expected, Internet Infrastructure exhibited high comovement in several months during this period and it had months with very high returns (mean monthly return through February 2000 =
### Table 4  Fixed-effects regression of comovement on return, January 1996 through September 2000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Unadjusted comovement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of stocks (per 100)</td>
<td>$-0.193 (0.032)$</td>
<td>$-0.183 (0.031)$</td>
<td>$-0.210 (0.031)$</td>
<td>$-0.208 (0.030)$</td>
</tr>
<tr>
<td>Mean return</td>
<td>$-0.227 (0.017)$</td>
<td>$-0.305 (0.017)$</td>
<td>$-0.248 (0.020)$</td>
<td></td>
</tr>
<tr>
<td>Mean return$^2$</td>
<td>$0.947 (0.074)$</td>
<td></td>
<td>$1.447 (0.108)$</td>
<td></td>
</tr>
<tr>
<td>Mean return$^3$</td>
<td></td>
<td></td>
<td>$-1.436 (0.226)$</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$0.236 (0.007)$</td>
<td>$0.238 (0.007)$</td>
<td>$0.237 (0.007)$</td>
<td>$0.232 (0.007)$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.442</td>
<td>0.469</td>
<td>0.491</td>
<td>0.497</td>
</tr>
</tbody>
</table>

(b) Adjusted comovement

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean return</td>
<td>$-0.334 (0.126)$</td>
<td>$-0.511 (0.135)$</td>
<td>$-0.518 (0.135)$</td>
<td></td>
</tr>
<tr>
<td>Mean return$^2$</td>
<td>$2.124 (0.575)$</td>
<td></td>
<td>$2.068 (0.842)$</td>
<td></td>
</tr>
<tr>
<td>Mean return$^3$</td>
<td></td>
<td></td>
<td>$0.160 (1.76)$</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$0.008 (0.011)$</td>
<td>$-0.008 (0.013)$</td>
<td>$-0.160 (0.013)$</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.502</td>
<td>0.504</td>
<td>0.504</td>
<td></td>
</tr>
</tbody>
</table>

n observations (industries) = 3813 (101). Standard errors in parentheses. All coefficients are significantly different from 0 at $P < 0.001$. Each model includes a dummy variable for each industry, excluding one industry as a reference category.

---

**Figure 8**  Predicted relationship between return and adjusted comovement (from Table 4b, model 2).
0.279) and several months with quite low returns (mean monthly return from March 2000 through September 2000 = –0.074). However, it is striking that the months with high comovement are precisely those during which Internet Infrastructure stocks showed the lowest returns. The contrast between December 1999, during which time Internet Infrastructure stocks achieved spectacularly high returns yet moderate to low comovement, and April 2000, when returns were very low and comovement was exceptionally high, puts the larger pattern in bold relief.

As indicated in Figure 11, the negative association between mean return and comovement is quite evident for Telecom Equipment/Wireline as well. In fact, these two categories had the most negative relationship between comovement and return ($r = -0.695$ and $r = -0.676$, respectively) of any category that had at least one year of data and they were followed by seven other high-tech industries. However, as indicated by our analyses in Table 3, these categories were merely extreme manifestations of the general pattern. Overall, 78% (77 of 98) categories exhibited a negative relationship between comovement and return. The median correlation was –0.19, which was slightly below the level observed for INM (–0.22).

### 4.2 Consistency in the comovement structure

Finally, we investigate an additional feature of comovement that has implications for the issue at hand. In particular, evidence as to whether or not investors were sufficiently discriminating in their approach to internet stocks may be found in the extent to which

---

18Note that observers typically date the ‘crash’ in internet infrastructure and telecom in the autumn of 2000.
the structure of comovement among member stocks was consistent from period to period. If investors apply a well-developed theory of value to the stocks in a given category, they should be consistent in the distinctions they make between such stocks. That is, they will continue to group certain pairs of stocks together while others will repeatedly be treated as different from one another. Various comovement structures are possible. For example, the distinctions made by investors may reflect a center–periphery
structure in that certain stocks are considered prototypical of the category and others more peripheral. Alternatively, the structure may comprise relatively distinct sets of subcategories such that there is high comovement within sets but low comovement between them. Regardless, it is possible to examine the extent to which the same structure is found across periods. To the extent that investors are consistent in the theories they use to group value the stocks in a given category, the pairs of stocks (hereafter, ‘dyads’) that exhibit high (low) comovement in one period should also exhibit high (low) comovement in subsequent periods. Conversely, there should be little consistency in the dyadic comovement structure if investors are not very discriminating among members of a category. At the extreme, where investors are simply purchasing or selling stocks on the basis of their category membership and without regard to individual differences, the comovement structure will be random in any one period and uncorrelated across periods. If internet stock investors employed such a crude strategy, internet stocks would exhibit very low comovement consistency.

We illustrate comovement consistency in the case of a relatively consistent category: Airlines. In Figure 12, the unit of analysis is the stock dyad. The horizontal axis indicates the dyadic correlation in daily return between each pair of stocks during April of 1998. Looking from right to left in the graph, we see that there appears to be a rough core-periphery structure to this category. In particular, the core group appears to consist of the major carriers (e.g. American, Delta and US Air) and the periphery consists of some smaller carriers (Comair, Reno) and a former major carrier (TWA) that are themselves treated as relatively heterogeneous. A mix of major carriers (Continental, Northwest and United) and mid size carriers (Alaska, America West and Southwest) seems to occupy a position just outside the core during this period. This overall structure is largely reproduced in the next period as reflected in the displayed regression equation. Thus, we see that the same core major carriers also display a high degree of comovement with other airlines in the next period and they are once again either orthogonal to or negatively correlated with the smaller carriers. Some differences across the two periods are notable as well. In particular, TWA is treated by market participants as more similar to other airlines stocks in May 1998 than in April 1998, while Reno and Comair are regarded as more different. Such change may be due to the arrival of information that is more idiosyncratic in one period and more general to the industry in another, or to change in the prevailing market opinion as to how much a stock is subject to general industry factors or firm-specific ones. The general picture that emerges in the case of Airlines is a fairly high consistency in the comovement structure, though investors clearly recognize important differences among such stocks.

How does this pattern of comovement consistency compare with internet stocks and with other categories? In Figure 13, we display the pattern for the INM category in the same period. The overall level of consistency is lower than in the case of Airlines, as would be expected if internet investors were indiscriminate in their purchases of such stocks. Yet, as the results from the regression equation suggest, it is hardly the case that the comovement structure in April 1998 is unrelated to that of May 1998. Indeed, the
Figure 12. Dyadic comovement in May 1998 by dyadic comovement in April 1998: Airlines. Regression equation ($R^2 = 0.265$): $\text{Comovement}_{\text{May98}} = 0.025 + 0.65 \times \text{Comovement}_{\text{Apr98}}$. Key: ACAI, Atlantic Coast; ALK, Alaska Air; AMR, American; AWA, America West; COMR, Comair; CAI, Continental; DAL, Delta; LUV, Southwest; NWAC, Northwest; RENO, Reno; TWA, TWA; U, US Air; UAL, United.
overall pattern has much in common with Airlines. In INM as well, there is a clear core, which consists of the ‘portals’ (Lycos, Excite, Infoseek and Yahoo!) as well as Amazon.com and some near-core stocks such as America Online and Doubleclick. There is also some variation between the two months in the extent to which some of the more peripheral stocks are treated as members of the category. For instance, Sterling Commerce is treated more as an INM stock in April than in May and the reverse is true for Cendant.

In Figure 14, we display the distribution of comovement consistency from January 1996 through August 2000. Each box plot portrays the distribution of comovement consistency across industries, where comovement consistency is defined as the correlation between a set of stock pairs’ dyadic comovement in month $m$ with their dyadic comovement in month $m+1$. The patterns reinforce our conclusions above. If low comovement consistency indicates that investors are not making meaningful distinctions among category members then it appears that investors in INM stocks were making such distinctions. While there are several months in late 1996 and early 1997 in which comovement consistency was low for INM stocks, there are many more months when it was relatively high. Moreover, comovement consistency is clearly lower during 2000 than during 1998 and 1999. This reinforces our finding that investors make fewer distinctions during periods when they are lowering their valuation for a category than vice versa.

A comparison of Figure 14 with Figures 6 and 7 suggests that the patterns that we observe for comovement consistency may reflect variation in the level of comovement. In particular, it appears that comovement consistency was high for INM stocks during the period when comovement itself was high. This pattern itself is awkward for those who might like to interpret high comovement among INM stocks as indicating that investors were not sufficiently discriminating. However, it is worth investigating the level of comovement consistency when controlling for the level of comovement in a category. We display the results of this analysis in Figure 16. The data points in this graph are the studentized residual from a monthly regression of comovement consistency from month $m$ to month $m+1$ on comovement in month $m$. And the patterns look very similar. While the level of adjusted comovement consistency in 1998 and the first half of 1999 is somewhat lower than that for comovement consistency, the difference is small. Overall, INM investors seem to have made distinctions among internet stocks at a level that was comparable to, and even exceeded, the level observed among other categories of stock. Moreover, the level of consistency drops precisely where prevailing interpretations of the internet phenomenon would presume that it would increase: in the last year, when investors were presumably becoming more discerning.

---


20It is intriguing that the drop in both comovement and comovement consistency occurs before the ‘crash’ in INM shares, which commenced in January 2000. This suggests that there may be some serial relationships between comovement and return in addition to the contemporaneous relationships we analyze in this paper. Future work should examine whether market movements may be predicted by prior patterns of comovement.
Figure 13  Dyadic comovement in May 1998 by dyadic comovement in April 1998: Internet & New Media. Regression equation ($R^2 = 0.097$): 

$\text{Comovement}_{\text{May98}} = 0.12 + 0.33 \times \text{Comovement}_{\text{April98}}$. Key: AMZN, Amazon.Com; AOL, America Online; ATHM, At Home; CD, Cendant; CNWK, CNET; DCLK, Doubleclick; HRBC, Harbinger; LCOS, Lycos; NEWZ, Newsedge; OMKT, Open Market; PTVL, Preview Travel; RNWK, Realnetworks; SEEK, Infoseek; SPLN, Sportsline USA; SE, Sterling Commerce; XCIT, Excite; YHOO, Yahoo!
Figure 14. Monthly distributions of comovement consistency. Mean correlation between dyadic comovement in month $t$ with month $t+1$. Outliers are black dots. Asterisks (linked) are mean coherence for Internet & New Media.
5. Discussion

In the foregoing analyses, we have used the issue of comovement as a lens through which to address the larger question of whether internet stocks were priced in a manner that departed significantly from the way other types of stocks are priced. Before considering how our results speak to this question, and the larger question of how stocks are valued, we highlight how two aspects of our analysis of comovement contribute to the understanding of how categorical systems channel market dynamics.

5.1 Method for comparative comovement analysis

First, we have introduced a novel method for obtaining comparable samples of assets from the categories that comprise a market, which facilitates analysis of variation across such categories. Consider the drawbacks of alternative approaches such as using the SIC/NAICS codes to divide securities or by using membership in indexes created by securities firms or media organizations. The former approach is problematic because the SIC/NAICS system is administered by the US government to satisfy a variety of purposes and constituencies and is updated relatively infrequently. The latter approach is particularly limited for comparing patterns across industries because different organizations are likely to apply different methods for assignment of a stock to a category. Moreover, the procedure for classifying a stock is typically unknown.

By contrast, the method we have developed in this paper exploits the implicit classifications made by the largest investors in their interactions with the key market intermediaries. When an institutional investor votes for an analyst as a top performer in a particular category, she implicitly identifies the analyst as a specialist in that category and indirectly assigns the stocks covered by the analyst to the category in question. By analyzing the full structure of such implicit classifications, we emerge with a set of comparable subsamples of stocks based on the views of the key actors in the market. As a result, we have been able to present what we believe are the first results on the distribution of comovement across the categories that comprise the market. Moreover, we have shown how this procedure may be used to identify the relative coherence of a category, something that is unknowable with alternative procedures. That is, rather than reifying a category and assuming that all classifications are equally informative to market participants, we may measure the extent of heterogeneity among the members of a category and we may use that information in analysis. Finally, note how our approach is distinct from analyses of comovement that focus on the influence of institutional categories such as the S&P 500 (Barberis et al., 2003). In such analyses, categorization represents an exogenous shock that is outside of investors’ control. Yet such analyses are unhelpful if we wish to understand how the processes by which market participants themselves interpret and classify securities, and the influence that such classifications have. Our approach is useful for this purpose because it emerges from the actions taken by key investors and intermediaries to order the universe of stocks they must confront.
Figure 15  Monthly distributions of adjusted comovement consistency. Studentized residual from regression of comovement consistency on comovement in month $t$. Outliers are black dots. Asterisks (linked) are mean coherence for Internet & New Media.
5.2 Comovement and price declines

A second notable feature of analysis concerns our finding that comovement (and comovement consistency) is highest during periods of price erosion. We know of no precedent for this result, which is unsurprising because it is unobtainable unless one employs a procedure such as ours for generating comparable subsamples of stocks. Why are members of a category treated as being more alike during periods when those stocks are being devalued than when they are increasing in value? It is tempting to interpret this pattern as being due to a kind of ‘threat-rigidity’ whereby the prospect of severe loss induces a restriction of information processing, such as a narrowing in the field of attention, a simplication of information codes, or a reduction in the number of channels used’ (Staw et al., 1981: 502). Yet the extension of the notion of threat-rigidity to securities markets is not straightforward. First, it is not clear whether our results suggest rigidity in the way that investors make sense of a category during periods of decline as much as an exaggeration of their general disposition towards member stocks. Second, while it may make sense to generalize experimental results on threat-rigidity from individuals to collective actors such as groups and organizations, the market cannot be seen as a collective actor. The challenge lies in uncovering how such dispositions aggregate to form patterns in price. Thus, we hope that our intriguing results will attract deeper inquiry as analysts move away from simply assuming that the market ‘knows’ how to interpret information and engage in direct investigations of how that interpretive process actually works.

5.3 The internet stock phenomenon: simply deluded

How do our findings contribute to our understanding of the internet stock phenomenon and to stock-market valuation more generally? As reviewed above, prevailing accounts suggest that the extraordinarily high prices for internet stocks were symptomatic of a breakdown of market efficiency. In particular, prices became decoupled from fundamentals either because: (i) speculators engaged in a self-fulfilling dynamic in which their expectations that prices for internet stocks would increase led to everincreasing valuations of internet shares; or (ii) constraints on arbitrage (short-selling) allowed the market to become dominated by unsophisticated investors. According to either account, investors were ‘exuberantly indiscriminate’ as they drove

---

21Put differently, why is that Tolstoy’s famous quip, ‘Happy families are all alike; every unhappy family is unhappy in its own way’ needs to be reversed when applied to stocks rather than families? Thanks toChip Heath for pointing out this irony.

22With these caveats in mind, we inquired of several psychologists if they knew of research that showed that individuals are less discriminating under situations of crisis or threat than they are in more favorable conditions. We were informed that, if anything, experimental studies tend to show the opposite. In particular, the ‘regulatory focus theory’ propounded by Higgins and colleagues suggests that, when subjects are induced to focus on goals of safety, they tend to be more conservative in their theorizing than when they are made to focus on heightening their sense of accomplishment (see Liberman et al., 2001). We thank Lorraine Idson for her guidance in this area.
up the value of internet stocks by focusing on them as members of an undifferentiated
category and without regard to the differences among them.

The evidence we have presented helps rule out interpretations that assign primary
importance in accounting for the internet stock phenomenon to a disregard for
differences between stocks. We have shown that the process by which internet stocks
were valued was actually not substantially different from that observed among other
categories of stock. In particular, we found that: (i) comovement among internet stocks
was moderate to high but did not attain the very high levels that are implied by
prevailing accounts; (ii) when high comovement is found, it is typically during periods
of value depreciation rather than appreciation; and (iii) investors in internet stocks
were fairly consistent in the pattern of distinctions that they made from period to
period. Overall, it is striking how similar are the comovement patterns for the internet
category to other categories.

How can we square the unjustifiably high prices commanded by internet stocks in
the late 1990s and early 2000 with the fact the processes by which investors priced these
stocks are largely indistinguishable from those that determine the prices for other
stocks? We believe that the alternative view we sketched above helps explain this
paradox. In particular, the internet stock phenomenon was a drawn-out period during
which a theory of value was adopted, diffused among investors but then was discovered
to be wrong. That this incorrect theory had sustained itself for as long as it did may be
due to the factors emphasized by prevailing accounts. That is, speculation that prices
would rise persisted because speculators recognized the hold the theory had on one
another. At the same time, one reason that more pessimistic, alternative theories did not
emerge to compete with the dominant, optimistic theory may have been the absence of
a vehicle for such pessimistic views to be expressed. Regardless of how long the theory
lasted or what sustained it, it is important to emphasize that it was indeed a theory
rather than mere irrationality.

The view that market pricing is governed by theories that are always provisional and
may sometimes be wrong is founded on the recognition that the market faces inherent
limitations in its capacity for closing the gap between price and value. As Zuckerman
(2004) has argued, it is reasonable to assume that no such gap will exist to the extent
that: (i) information is widely available; and (ii) it is widely recognized how one should

23Thanks to an anonymous reviewer for phrasing our contribution in this way.

24Of course, the market’s opinion of the internet may yet reverse, though this seems extremely unlikely
at this point. If such a reversal did transpire, it would imply a spasmodic learning process that is hard to
square with orthodox theory.

25That is, short-sellers were constrained from full participation. Note also that, even in the absence of
short-selling there is a crucial check on the price level when prices get too low but not when they get too
high. In particular, to the extent an asset enjoys an income stream that is greater than its market price,
an investor can potentially buy the asset (e.g. take a public firm private). However, there is no
comparable, non-speculative vehicle for disagreeing with market prices when they get too high.
interpret that information. However, it is erroneous to assume, as does the efficient markets hypothesis, that the problem of interpretation is solved by making information available (Dreman, 1977). This assumption is based on a model of the market as a controlled environment for learning, in which market participants come to understand the implications of a particular type of economic information as the implications of that information are observed in multiple trials in the same conditions (Brav and Heaton, 2002). Yet the market is anything but a closed laboratory because the economy itself is exposed to continuous exogenous shock and internal change that cast doubt on prevailing models (Morris, 1996: 1128; Zuckerman, 1999: 1411; cf. Winter, 1986; Romer, 1993; Spotton, 1997; Shiller, 2000). This does not mean that the market does not tend towards a common interpretation of information about an asset. But such convergence should vary and it should take time for it to occur (Zuckerman, 2004). In the meantime, the market may be dominated by theories of value that, in hindsight, turn out to have been incorrect.

This view of the market has important implications for managers and scholars. From the perspective of scholars, our results underline that it is important to complete the dismantling of the efficient markets hypothesis but to do so in a way that preserves its core insights. Market prices should indeed be understood as disciplined by the practice of arbitrage: the arrival of information is incorporated into prices almost immediately as investors apply prevailing theories of value to interpret that information and derive its implications for prices. Yet such interpretations may be wrong in the short and even intermediate term because market participants may not have had the opportunity to observe such news and its implications on enough occasions and in the right conditions. As such, both managers and scholars should take current prices with a grain of salt. When the market responds to a strategic or operational change, it does so as interpreted by the prevailing theory of value. The resulting change in price is thus only as good as the theory in question. In most cases, it makes sense to believe that the prevailing theory is relatively accurate. In others, it does not. As a result, managers would do well to keep their own counsel rather than manage for market opinion (though at the same time to try to understand the theories that are in use by market participants). Similarly, scholars would do well to retire the ‘event study’ from their analytic arsenal, except in limited cases. In particular, changes in market price that occur as a result of an event should be understood as indicating the application of a possibly erroneous interpretation to that change rather as the ‘true’ meaning of that event, as is assumed by countless event studies found in the finance, economics and management literatures (but see e.g. Westphal and Zajac, 1998).

6. Concluding remarks

In concluding, we note broader implications of our approach for the emerging stream of research at the intersection of organization ecology and network theory. The first point of interchange between these research traditions occurred in the 1980s and 1990s
around models of competition and, in particular, the related concepts of niche width and structural equivalence (see Freeman and Hannan, 1983; McPherson, 1983; DiMaggio, 1986; McPherson et al., 1992; Burt and Talmud, 1993; Baum and Singh, 1994; Podolny and McPherson, 1995; Podolny et al., 1996). More generally, the basis for theoretical and empirical convergence derived from a common interest in the patterns of resource flows upon which firms and other social actors depend. More recently, concurrent developments in the two research traditions reflect a growing interest in what Podolny (2001) has labeled structures as ‘prisms.’ For network analysts, such a reorientation involves analysis of patterns of valuation or appraisal, which generate vertical (e.g. Podolny, 1993; Benjamin and Podolny, 1999; Stuart et al., 1999) or horizontal (Zuckerman, 1999, 2000; Zuckerman et al., 2003) differentiation, or some combination of the two (e.g. Park and Podolny, 2000; Philips and Zuckerman, 2001; Zuckerman and Kim, 2003). For ecologists, the recent shift centers around the question of how organization forms are defined and, in particular, how key audiences establish rules for delimiting organizational identity in a way that constrains organizational life chances (see Carroll and Swaminathan, 2000; Swaminathan, 2001; Pólos et al., 2002).

It is important to note what is common to both research streams. First, both ecologists and network analysts share an interest in the audience and how it constructs and applies its valuation models. Second, each upholds the distinctively sociological premise that a basis for differentiation is socially constructed in the process of collective valuation, which then governs dynamics in a way that cannot be fully reduced to initial tastes and endowments, or what might be termed ‘fundamentals’. Indeed, if location in a network of appraisal, or if identity with respect to an audience’s mapping rule, simply reflects such fundamentals, then one need not look beyond them to understand behavior and outcomes of interest. This issue is particularly salient since network analysts and ecologists have increasingly come to focus on market contexts where the penalties for deviations from fundamentals are presumably quite strong.

Our paper makes progress on a question that lies in this new point of intersection between ecology and network analysis. In particular, we shed light on the processes by which market audiences produce horizontal differentiation by grouping like and separating unlike. We study these dynamics in the stock market, which represents an excellent context due to two properties particular to financial markets. First, whereas patterns of valuation or interpretation are invisible or only periodically observable in most domains, capital markets require the continuous enactment of interpretive schemes through reaction to new information. Each slice of time affords a window into how the marginal market participant categorizes assets at that moment. The marginal valuations, and hence the prices, of securities that are regarded as similar move in the same direction in response to material news. Second, not only are classification patterns continuously observable; financial markets also afford a fundamental baseline against which to assess the impact of such patterns. Thus, we may investigate whether the observed patterns of differentiation indeed depart from what may be expected due to the assets’ true values and we may then begin to inquire how such departures occur.
Indeed, our ability to compare patterns with such a fundamental baseline leads to a conclusion that runs somewhat against the general thrust of the early work on interpretive schemes and codes. In particular, we have shown evidence that such structures may be considerably more fluid than is commonly supposed. This in turn suggests new paths of inquiry: rather than simply assuming them as hard constraints, our results call for research into the processes by which old theories of valuation are overturned and new ones emerge and, in what circumstances, such dynamics crystallize into relatively stable interpretive schemes that constrain strategy.

Acknowledgements
We would like to thank audiences at the Economic Sociology seminar at MIT and the Organizational Behavior seminar at Harvard Business School for their feedback on early versions of this work. We also appreciate the comments of participants at the NYU Stern Organizational Behavior, the Harvard Workshop in Applied Statistics, and the Stanford 25th Anniversary Conference on Organizational Ecology. Finally, we are grateful for the assistance of Denise Murrell and Diana Blaney of Institutional Investor magazine. We are solely responsible for any remaining errors.

Address for correspondence
Ezra W. Zuckerman, MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142, USA. Email: ewzucker@mit.edu.

References
Barberis, N., A. Shleifer and J. Wurgler (2003), ‘Comovement,’ working paper, University of Chicago Graduate School of Business.


Shivaram, R., M. Venkatachalam and S. Kotha (2002), 'Managerial actions, stock returns, and