Columbia University
Graduate School of Business
B7313
Securities Analysis

EMBA – Elective C
Fall 2004
Office Hours: By appointment and by telephone.

Professor Paul Johnson
Room 105 Warren Hall
Office Phone: (212) 293-3402

Academy Awards, Sharon Stone and Market Efficiency

Market efficiency may be the most emotional topic that market practitioners can discuss.

Virtually all of the academic studies suggest that the efficiency of the market is remarkably high. In fact, market anomalies are rare and usually apply to only a handful of specific stocks for a specific period of time. Despite these studies, most, if not virtually all, market practitioners believe that the efficiency of the stock market lies somewhere between nonexistent and inconsequential. We find this paradox to be interesting.

Academy Awards

We have had the opportunity to conduct the follow experiment. We asked students in a investment class to pick individually their best selection as to who would win the Academy Awards. We picked 12 categories for this contest; the first six categories were the most obvious[1] and the second six were somewhat obscure[2], at least to people outside Hollywood.

To make the contest interesting, the student with the greatest number of correct answers would win $50, depending on the contest. Being starving graduate students, it was more than enough money to keep them interested and motivate them to pick their "best" guesses. In addition to tallying each student's entrant, we also tallied how many student's selected each nominee within each category (median). We then determined that each nominee with the greatest number of votes in each category to be the "group's selection" and we considered this to be the consensus view. We then waited for the actual results to be announced, usually a few weeks later. The results always surprised us.

We have conducted this experiment five times. The results from each year are presented in Tables 1- 4 at the back of this essay. Table 1 presents the results from 1995 and 1996, Table 2 present the 1997 results, Table 3 present the results from 2002 (we took a few years off from the

[1] Best Actor Male, Best Actor Female, Best Supporting Actor Male, Best Supporting Actor Female, Best Picture, and Best Director
[2] Original Screenplay, Adapted Screenplay, Cinematography, Film Editing, Original Score and Original Song

contest), and Table 4 the results from 2004 (again we skipped a year—hey, it is a part-time gig!). In 1997 we conducted the experiment in three different investment classes, two at the Graduate School of Business, Columbia University, the third at the Graduate School of Business, University of California, Berkeley. The 1997 results included the largest sample size and resulted in the most robust outcome, all of which is summarized in Table 6. In 2002, we conducted the contest in two different investment classes at Columbia, with a bit smaller population than the 1997 results. The 2002 contest was the most surprising because it produced the least robust results of all of the experiments. Although the size of the participants in the 2004 was small (at least by recent standards in the contest), the results mimicked the outcomes of prior years and reversed the unusual results produced in 2002

What is interesting in the results from the contest in each of the years is that the group selection, the "consensus view," has an amazing track record at picking the winners of the actual Academy Awards. In fact, although the consensus view was beat for the first time in the 2002 contest, it had never lost to a student in any other year. The consensus had tied the best student in some of the years, but had an almost perfect record in predicting the winners. The second surprise from the contest is that the consensus did remarkably well despite the low score of the average student (this held true even in 2002). In each case, the consensus view scored more than the correct answers of the average student—in 1995, 1996 and 1997 the consensus scored more than twice the average student. This is fascinating. The third surprise is how few data points are needed before a powerful consensus is created. For instance, there were only 22 entrants in the 1995 contest and, despite the low population, the results are equally robust as contests in the later years based on a higher number of entrants. Although the data suggests that the contest does not need many contestants to produce a robust outcome, the 1997 contest shows what happens as the population increases. Using the summary data presented in Table 2, the results show that the consensus managed to "guess" 11 correct answers out of a possible 12 that year. The best human scored 10 and the average contestant scored only 4.5 correct answers. In 2004 the consensus correctly guessed all 12 winners for the first time in the contest. Although 2004 is perceived as an "easy" year, with all of the favorites pulling through to win their category, the average student guessed only 7 correctly. Although the average for correct answers from the students was well above past averages, it was still below the results for the consensus. In addition, although one student also guessed all 12 categories correctly, thus matching the consensus, the second highest scoring student guessed "only" 11 correctly, matching the pattern in prior contests. Therefore, despite the relatively higher scores in 2004 (hence the moniker of being "easy"), the results followed exactly the same pattern of past contests.

As a consequence of the results from all of the contests, it is clear that the consensus is a powerful force.

Consensus Mechanics - The importance of being independent

What is going on here? How can a consensus produce such amazing results, when the average participant scored so poorly?

The key to understanding the results is the notion of independent errors. The explanation goes like this. Each participant will make mistakes as to their individual guesses. But few will guess all 12 categories incorrectly. Most will have an informed opinion on at least a few of the categories. Each person will have seen the movies or read a review or two; they will be exposed to public opinion; and some will be influenced by the winners of the Golden Globe Awards (that were announced earlier in the year). Nevertheless, each person will have a different opinion as to the potential winners. The interesting part of the contest is that the errors begin to cancel each other out. There is only one actual winner (unknown at the time), but there are four non-winners (although even being nominated is a form of winning!). Therefore, simple statistics suggest that the winning guesses will begin to reinforce each other, while the losing guesses will begin to cancel each other out. However, all of this presupposes that the guesses are not completely random. If each student guessed randomly, then the statistics would predict that each of the five nominees would receive the same number of votes and the no real consensus would be formed. Therefore, the contest assumes that each student has at least a non-random opinion (technically, we only need some of the students to have a non-random opinion for the consensus to form, but the consensus becomes more robust as the "non-randomness" of opinion increases.) The key to the contest is that the various opinions begin to cancel each other out when they are wrong and reinforce each other when they are right. Therefore, all that is needed is a slight bit of "correct" information to enter the contest for a consensus view to begin to emerge. We call this the power of independent errors and very powerful it is. But the contest does not always work this way.

Repeatability - one may win this year, but will not the next

Despite the apparent power of consensus, the systems is not without its flaws and potential weaknesses. What happens when errors no longer remain independent? Interestingly, the student's guesses are influenced by many sources, but sometimes the world decides that one particular entrant is "just going to win." This was the case for the English Patient as the Best Picture nominee in 1997. The movie had received the most nominations, had won the Golden Globe Award and had received the critics support. It was just going to win! The popular opinion was reflected in the contest results. Out of a possible 128 entrants, 100 selected the English Patient to win—84% of the student selections. Not surprising, the movie did win. Chalk one up for the consensus, and most of the students.

But sometimes the popular press is not correct in "pre-selecting" a winner as was the case for Best Actress in 1996. In a similar fashion as in the Best Picture contest in 1997, the press decided that Sharon Stone was going to win for her role in Casino. She had won the Golden Globe Award and the hearts of the critics. Once again this opinion showed itself in the contest results. Roughly 85% of the students selected her to win. Nevertheless, she was edged out by Susan Surandon for her role in Dead Man Walking. In this case the consensus was wrong.

Both situations represent a case of non-independent errors. Many of the students were influenced by a similar, external opinion. Despite strong evidence that no one "knows" who will win the award, the students were overly influenced by the opinion of the popular press. This resulted in non-independent thinking and a fragile consensus. In one case, the consensus was right, but in the other it was wrong. Independent of the outcome, both cases are examples where

the independent thinking collapsed and produced a fragile consensus. Interestingly, a similar phenomenon of non-independent thinking occurs when group members are allowed to discuss their guesses ahead of time. In these situations, the loudest, and often the most confident, in the group tends to have influence over the group's thinking and can influence individual selection. Although we have not run the experiment with a group discussion, the academic literature on management theory is chock full of similar experiments. This phenomenon is referred to as "group think," and can be viewed simply as a case of non-independent thinking or non-independent errors. This example parallels the role of the press or a well-respected analyst as a lighting rod to galvanize opinion in the stock market, many times resulting in a less robust consensus view!

One additional potential flaw of the system of selecting a consensus is when there are too few contestants. It is clear from our work that the more, the merrier. The power of the consensus clearly strengthens as the number of contestants increases (although, as we pointed out earlier, a key criteria is that the individuals are at least partially independent in their thinking). How many are needed is beyond the scope of our experiment, but it is clear that as the number of participants decreases, the power of the consensus weakens.

One of the key outcomes of this experiment is that when "enough," at least partially informed, and reasonably independent participants, are allowed to select individual results, the consensus results will be powerful. In most cases, the consensus will be more accurate than all of the participants and significantly more robust that the average participant. This is a very powerful metaphor for the stock market.

It is also important to note that it would be unlikely that a single individual would do well each year. One student may produce the highest number of correct answers one year, maybe even beating the consensus that year (it has only happened once in the five contests), but it is doubtful that the same student would do it a second year. In most cases, the student was lucky. Odds are high that the luck would not return to the same student two years in a row.

Is there an Academy Award in the Stock Market?

We believe that there are many parallels between our Academy Awards contest and the stock market. In the stock market, individual compete with each other to have a better guess as to the true value of individual securities. No participants knows for sure the true value of these stocks and no participant has complete knowledge. Most believe themselves to be well informed. Nevertheless, we all make mistakes. But rarely are these mistake the same across all individuals. In fact, many of our best guesses have a fare amount of correct information. The things that we know to be true tend to be common among other participants and the things that we "know," but are not correct (errors), tend to be individual. Therefore, our errors tend to cancel themselves out while our correct guesses tend to reinforce themselves. In this same way, the Academy Awards contest draws an interesting and powerful metaphor for investing.

Interestingly, the metaphor actually maps well even when it breaks down. For instance, there are times when public opinion is influenced by external influences or information. In some of the

instances this information is powerful enough to influence individual decision making. In these cases, the errors become non-independent because we are all starting to think alike. This is a form of "group think." In the stock market we tend to call these manias or panics. Manias when the "consensus opinion" of value becomes overly inflated and panics when the "consensus opinion" becomes too negative (actually, unlike the Academy Awards contest, the stock market is dynamic and manias/panics generally begin to feed on themselves exaggerating the direction of the "consensus view"). Manias, and panics, are merely a case of group think and non-independent errors. Interestingly, the simple model of guessing the Academy Awards can explain much of the stock market behavior in these situations.

The stock market also has periods of low activity, when there are not enough participants in a situation to produce a robust consensus. We tend to call these situations neglected stocks. Once again, the Academy Awards contest is an apt metaphor.

Although the 2004 contest put the results back on track, we continue to be troubled by the results in 2002? Unlike all other years, the result for 2002 did not follow the same pattern and the consensus did not produce the same impressive results. Interestingly, although the best human guesses did in fact beat the consensus for the first time, the number of correct guesses by the winner was no better than in years past and the average student was well below prior year contests as well. It may be important to note that one distinct feature of the 2002 contest is that no clear consensus "opinion" emerge in many of the categories and, even the Hollywood experts agreed, the Oscar winners that year were "surprising." Therefore, the lesson to be learned is that when no clear consensus is at work, surprise outcomes produce unusual winners. Ironically, the 2002 Academy Awards mirror the U.S. stock market for the same period: no clear consensus on market direction, unusual outcomes in stock performance and unpredicted winners (stocks and money managers).

Market Efficiency

How does all of this relate to market efficiency? Even the academics have some trouble in defining precisely what is market efficiency. Nevertheless, we will settle on a simple definition. We believe that a market is efficient when no individual can out guess the market as to the true value of any individual stock consistently over time. On any given day, we may concede that a few individuals can out guess the market as to the true value of a specific stock (ala the 2002 Academy Awards), but these same individuals cannot out guess the market as to the true value of this same stock consistently over time. Another way of saying this is that no individual can out guess the market as to the value of multiple stocks on a consistent basis.

The Academy Awards contest certainly supports the notion that the consensus is a very powerful mechanism, particularly when a few assumptions are made. What is interesting is that nowhere in the contest have we asked the participants to be "rational." Our only request is that they are independent from each other and at least partially informed. Once these two conditions are met, the consensus appears to be the most robust prediction as to the best guess or most true value at least over time (although 2002-like outcomes do happen in the stock market as well). I would suggest that for most stocks in the stock market, these conditions hold, most of the time. Yes, it

does appear that manias and panics can develop, but they appear to be relatively infrequent. Making the case that most investors believe themselves to be at least partially informed also seems to be an easy assumption. Finally, in terms of acting independently, again this appears to be a reasonable assumption, most of the time. With all of these conditions in place, the Acadeny Awards metaphor would suggest that the market is efficient for most companies, most of the time.

Where does this leave us?

Interesting question. When there are enough, informed investors acting independently, the consensus view will be more powerful than any individual's view, and the market will be essentially efficient for that security (remember that this has nothing to do with acting rationally). It is only when these conditions fail that the market becomes less efficient. When there are not enough participants, the stock becomes neglected and may become miss-priced. When investors stop acting independently, they may begin to exhibit group think and either a mania (positive) or a panic (negative) may ensue. The rest of the time, the market will be better than man, at least individual man.

So the next time you watch the Academy Awards, think about your investment strategy. And visa versa. Each time you invest, remember the power of the consensus.

Academy Awards Experiment

Table 1

	1995				1996			
	Consensus	Best	Humans 2nd Best	Average	Consensus	Best	Humans 2nd Best	Average
I	4	5	2	2.40	3	4	4	2.00
II	5	4	6	2.73	5	4	3	2.58
Total	**9**	**9**	**8**	**5.13**	**8**	**8**	**7**	**4.58**
Participants				22				33

Table 2

	1997 PJ				1997 MM			
	Consensus	Best	Humans 2nd Best	Average	Consensus	Best	Humans 2nd Best	Average
I	5	3	5	2.57	5	4	5	2.60
II	5	5	2	2.30	4	4	3	2.22
Total	**10**	**8**	**7**	**4.87**	**9**	**8**	**8**	**4.82**
Participants				28				55

	1997 PS				1997 All			
	Consensus	Best	Humans 2nd Best	Average	Consensus	Best	Humans 2nd Best	Average
I	5	5	4	2.48	6	5	5	2.55
II	5	4	4	2.33	5	4	3	2.27
Total	**10**	**9**	**8**	**4.81**	**11**	**9**	**8**	**4.82**
Participants				42				125

Table 3

	2002 PJ				2002 MM			
	Consensus	Best	Humans 2nd Best	Average	Consensus	Best	Humans 2nd Best	Average
I	2	2	2	1.70	2	5	5	2.10
II	2	4	4	1.70	3	3	3	1.70
Total	**4**	**6**	**6**	**3.40**	**5**	**8**	**8**	**3.80**
Participants				49				54

	2002 All			
	Consensus	Best	Humans 2nd Best	Average
I	2	5	5	1.90
II	3	3	3	1.70
Total	**5**	**8**	**8**	**3.60**
Participants				103

Table 4

	2004			
	MM	Humans		
	Consensus	Best	2nd Best	Average
I	6	6	6	4.00
II	6	6	5	3.00
Total	**12**	**12**	**11**	**7.00**
Participants				46