Just Lucky?
A Statistical Test for Option Backdating

Richard E. Goldberg[†]
James A. Read, Jr.[‡]

The Brattle Group

Abstract

The literature in financial economics provides convincing evidence of retroactive timing of executive stock option grants (option "backdating"). However, the literature does not yet contain a methodology for detecting backdating at individual companies. The present paper seeks to fill this gap. Specifically, it describes a rigorous statistical test for backdating based on publicly available data. In addition, it identifies a flaw in the methodology employed by the *Wall Street Journal* (Forelle and Bandler, 2006) to calculate the odds that the timing of executive stock option grants was purely random, and it provides experimental evidence that this flaw could in practice distort the assessment of odds.

†   Richard E. Goldberg, The Brattle Group, 333 Sacramento Street, Suite 1140, San Francisco, CA 94111, Richard.Goldberg@brattle.com.

‡   James A. Read, Jr., The Brattle Group, 44 Brattle Street, 3rd Floor, Cambridge, MA 02138, James.Read@brattle.com.

Just Lucky?
A Statistical Test for Option Backdating

Richard E. Goldberg
James A. Read, Jr.

The Brattle Group

1. Introduction

Option backdating—the retroactive selection of grant dates for executive stock options in order to set favorable strike prices—is currently big news, with over 100 companies under investigation by the Securities and Exchange Commission (SEC).[1] Regulatory scrutiny of the timing of executive stock option grants was initially stimulated by a line of academic research that goes back to Yermack (1997). In a sample of approximately 600 stock option grants from 1992 to 1994, Yermack found that the returns on the underlying stocks were normal in the period prior to the option grants but were abnormal (positive) in the period following option grants. After considering several possible explanations for these results, Yermack concluded that managers influenced compensation committees to award more options when they anticipated improvements in corporate performance. Lie (2005) examined a much larger sample of stock option grants from 1992 to 2002.[2] Like Yermack, he found abnormal returns in the period following option grants, but he also found abnormal (negative) returns in the period preceding option grants. In addition, Lie suggested a novel explanation for these findings, namely, that grant dates might have been selected *retroactively*, in light of observed (rather than anticipated) increases in underlying stock prices. Recently, Lie and Heron (forthcoming) report that evidence of abnormal returns around option grant dates declined sharply following the Sarbanes-Oxley Act, which mandated that grants be reported within two business days of the grant date. Thus, the academic research provides convincing evidence that backdating has indeed occurred.

The academic literature to date has addressed option backdating in the aggregate but not at the level of individual companies. However, earlier this year the *Wall Street Journal* (Forelle and Bandler, 2006) identified several companies for which there is a striking pattern of increases in the prices of underlying stocks immediately following option grants. The *Wall Street Journal* (*WSJ*) also reported eye-popping odds against these results having occurred by chance. For example, they reported that the odds that the returns on shares held by Jeffrey Rich, then CEO of Affiliated Computer Services, occurred by chance were only one in 300 billion; the odds of returns on shares held by Louis Tomasetta, President and CEO of Vitesse Semiconductor, were only one in 26 billion; and the odds of returns on shares held by Kobi Alexander, Chairman and CEO of Comverse Technology, were only one in 6 billion. The *WSJ* computed these odds using

---

[1] See, *e.g.*, the article by Pender (2006) in the *San Francisco Chronicle*.
[2] Other papers in this line of academic research include Aboody and Kasznik (2000) and Chauvin and Shenoy (2001).

an unpublished methodology that it devised and implemented with the assistance of a statistician. (The methodology is described in the same article.)

The purpose of the present paper is threefold. First, we identify a flaw in the statistical test the *WSJ* applied to detect option backdating. Second, we describe an experiment we performed using Monte Carlo simulation, which shows that this flaw could distort the assessment of odds in practice. Finally, we develop an alternative statistical test that does not suffer from the same flaw. To summarize briefly, the flaw in the *WSJ* methodology is related to the fact that it is based on annual rankings of returns. The alternative test we propose, in contrast, is based on returns but not on return ranks.

2. The WSJ Calculations

The results reported by the *WSJ* were obtained by applying the following methodology.[3] For each trading day in which there was a stock option grant for a particular company, the *WSJ* computed the 20-day return on the underlying stock.[4] It then ranked the return on each grant date in relation to the 20-day returns on all other days in the same year. For example, if the 20-day return on a grant date was the highest in the year, its rank would be 1, if the 20-day return on a grant date was the second highest in the year, then the rank would be 2, and so on. Then, under the assumption that each day's rank was random, they computed the likelihood of observing a pattern of return ranks as extreme as the ranks observed for the grant dates.

To illustrate the *WSJ* methodology, we will describe the result cited above for Jeffrey Rich of Affiliated Computer Services (ACS). Of the six grants received by Mr. Rich from 1995 through 2002, the returns for two grant dates were ranked highest in their respective years, the returns for two grant dates were ranked second highest, the return for one grant date was ranked third highest, and the return for one grant date was ranked fourth highest (see Table 1). We will represent this (unordered) outcome as {1,1,2,2,3,4}. If there are 252 trading days in the year, then picking one day randomly in each of six years (assuming no ties) has 252^6 (approximately 256 trillion) different possible outcomes. We can then count how many of these outcomes are at least as extreme as the outcome for the six ACS option grants. There are 19 different unordered outcomes at least as extreme but each of these outcomes can have permutations with the same unordered ranking. For example, in an outcome consisting of five highest ranked days and one second highest ranked day, the second highest ranking day could have occurred in any one of six years.

Table 2 lists the 19 different outcomes at least as extreme as the {1,1,2,2,3,4} outcome observed for the ACS grant dates. It also reports the number of equivalent ordered outcomes. The result is that only 864 of the roughly 250 trillion possible outcomes are as extreme as the outcomes associated with the ACS grant dates. This is equivalent to approximately one in 300 billion (= 256 trillion/864).

---

[3] See Forelle (2006).

[4] Lie's analysis suggests that most of the abnormal returns following stock option grants occurs within 20 to 30 days following the grant date, so although the *WSJ*'s choice of 20 days to compute returns was somewhat arbitrary, it is not inconsistent with the academic research.

To summarize, the *WSJ* methodology is based on the likelihood of randomly selecting days on which the 20-day return ranked as high as the 20-day return observed for the stock option grants to company executives.

3. A Problem with the WSJ Calculations

A key assumption in the *WSJ* calculations is that the selection of stock option grant dates is independent of the annual ranking of the 20-day returns. Clearly a company's price history is known at the time stock options are granted. This means that the returns are known and could be taken into account in a decision as to when to grant options. However, these returns also enter into the calculation of annual rankings. If the rankings are affected by information that is used in selecting the grant date, then the assumption that the grant dates are a random sampling of the days of the year is invalid.

Is this a minor technical issue or could it be significant? To explore this question, we conducted a two-part experiment using Monte Carlo simulation. In the first part of our experiment, we drew a large number of one-year sample paths of daily stock prices under the assumption that the price P_t follows a lognormal random walk, i.e., $\ln(P_{i+\Delta t}) - \ln(P_i) = \mu \Delta t + \sigma Z_i \sqrt{\Delta t}$, where the Z_i terms are drawn from a standard normal distribution and $\Delta t = 1$ trading day $= \frac{1}{252}$ year. To be specific, we assumed that $\mu = 20\%$/year and $\sigma = 50\%$/year$^{1/2}$. We then randomly selected a grant date for each sample path and computed the 20-day grant-date return. Finally, we repeated the simulation 100,000 times with the following result: The grant date had the highest ranked 20-day return in 385 of the 100,000 simulations. This is equal to 1 out of every 260 simulations, close to the predicted result of 1 out of every 252.

In the second step of our experiment we selected grant dates using an alternative rule: Offer the grant on the first day the price had dropped each of the prior ten days or, if that never occurred, on the last day of the year. After executing another 100,000 simulations (with the same random draws) using the alternative rule, we found that the grants were made on the highest ranked day 977 times out of 100,000. This is equal to 1 out of 102 days, about twice as often as predicted based on the assumption that the grant date was selected randomly. (Although the exact numbers change in each set of simulations, we consistently found that grant date selection using our simple strategy roughly doubled the chances of the grant occurring on the highest ranked days.)

If very simple strategies based on prior returns can double the likelihood of achieving highly ranked days, then strategies that company insiders can think of should be able to achieve even greater improvements in the likelihood of highly ranked grant dates. Therefore, a statistical test based on the assumption that the rankings of grant dates are independent of date selection cannot be trusted. For this reason, we conclude that the statistical test used for screening purposes by the *WSJ* is flawed.

4. An Alternative Test

The basis for the alternative statistical test we propose is the fact that, while annual rankings of returns may depend on past prices, the returns themselves should not. For example, in our simulations we found that the mean 20-day log-price return (defined as the difference in the natural logarithm of prices) following grant dates was almost identical for both the random and strategically selected grant dates: 1.65% for the randomly selected dates and 1.58% for the dates selected strategically. (The population return of 20%/year corresponds to 1.59% over 20 days and the standard deviation of estimates over 100,000 samples should be equal to .045%.) Therefore, a statistical test should be based on a comparison of returns around grant dates to typical returns, not on the annual rank of grant-date returns.

To follow up on this insight, we construct a T test for the difference in mean returns over grant dates and mean returns over typical trading days. We assume that the daily log-price returns $\ln(P_{i+1}/P_i) \approx (P_{i+1} - P_i)/P_i$ for trading days in the period $i=1,\ldots,N+W$ can be modeled as independent and identically distributed random variables of the form

$$\widetilde{X}_i \equiv \ln\left(P_{i+1}\Big/P_i\right) = \mu\Delta t + \sigma\widetilde{Z}_i\sqrt{\Delta t}$$

where the Z_i are independent standard normal draws. In this case, the W-day log-price returns for trading days $i=1,\ldots,N$ can be written as

$$\widetilde{Y}_i \equiv \ln\left(P_{i+W}\Big/P_i\right) = \sum_{k=0}^{W-1}\widetilde{X}_{i+k} = \mu W\Delta t + \sigma\sqrt{W\Delta t}\left(\frac{1}{\sqrt{W\Delta t}}\sum_{k=0}^{W-1}\widetilde{Z}_{i+k}\right) \equiv \mu_Y + \sigma_Y\widetilde{Z}_i^Y$$

where the Z^Y_i are standard normal draws that are correlated when within W days of one another such that

$$E\{\widetilde{Z}_i^Y \widetilde{Z}_j^Y\} = \begin{cases} \dfrac{W - |i-j|}{W} & \text{if } |i-j| \le W \\ 0 & \text{otherwise} \end{cases}$$

Our null hypothesis is that the set of n option grant dates were selected in a way that did not depend on the returns following the grant dates. If that hypothesis is correct, then the mean W-day log-price return on the grant dates m_n is an unbiased estimator of the population return μ_Y. Suppose then that we observe that m_n is much higher than the estimator m_N computed using the full data set of N trading days. We would like to develop a statistical test to see if such an observed difference is solely due to chance.

A T test can be used to test the significance of a difference in means. However, it requires that the trading data being compared with the grant-date data satisfy certain conditions, namely, that the W-day returns for each sample of trading data used to estimate the population mean and variance be independent—independent of each other and independent of the returns on each grant date. This can be accomplished by parsing the full set of trading data into a mesh of M

data points that are at least W days apart from each other and also from each grant date. (Since the data points are W-day mean returns, all data points less than W days apart will be correlated with each other.) We can calculate the mean and standard deviation of W-day returns over the mesh of parsed data points (m_M and S_M, respectively) and then perform a T test to see if an observation that m_n is much greater than m_M is statistically significant. The T test statistic will have M-1 degrees of freedom and be equal to

$$\tilde{T}_{M-1} \equiv \frac{(\tilde{m}_n - \tilde{m}_M)}{\tilde{S}_M / \sqrt{n}} \left(\sqrt{1 + \frac{n}{M} + \frac{2}{n} \sum_{\text{nearby grants } j,k} MAX\left\{0, \frac{W - |i_G(j) - i_G(k)|}{W}\right\}} \right)^{-1}$$

where

$$\tilde{m}_n = \frac{1}{n} \sum_{j=1}^{n} \tilde{Y}_{i_G(j)}$$

$$\tilde{m}_M = \frac{1}{M} \sum_{m=1}^{M} \tilde{Y}_{i_M(m)}$$

are the estimated means over the grant dates and over the parsed data points, respectively, where $i_G(j)$ and $i_M(m)$ enumerate the trading days corresponding to the n grant dates and to the M parsed data points, respectively;

$$\tilde{S}_M = \sqrt{\frac{1}{M-1} \sum_{m=1}^{M} \left(\tilde{Y}_{i_M(m)} - \tilde{m}_M\right)^2}$$

is the standard deviation of W-day returns estimated using the parsed data set; and the sum over "nearby grants" is a sum over distinct pairs of grant dates that are within W days of each other.

The observation that returns are higher following grant dates than they are on other trading days corresponds to a high T test statistic. Whether or not the test statistic is higher than would likely result solely from random chance can be determined by reference to values of the one-tailed T distribution with M-1 degrees of freedom. If the likelihood of achieving the observed test statistic is less than a reasonable threshold, the default hypothesis that option grant dates are selected without foreknowledge of future prices will need to be rejected.

5. Comparison of Test Results

To compare our test with the *WSJ* test, we computed the odds of achieving the results the *WSJ* reported for ACS, Vitesse Semiconductor, and Comverse Technology (see Table 3). In each case we verified the result reported by the *WSJ* based on our understanding of their methodology and then applied the T test described herein. Since for 20-day returns there are 20 different parsed data sets that our method would allow for the test (depending on how many days from the first day in the period we start our mesh of points), all of the results we report are from the parsed data set with the highest odds of occurring. Grant date information was obtained from

proxy statements filed by the companies. Stock prices were obtained from Bloomberg. The *WSJ* calculations addressed six CEO grants over the period 1995 to 2002 for ACS, nine CEO grants over the period 1994 to 2001 for Vitesse Semiconductor, and eight CEO grants over the period 1994 to 2001 for Comverse Technology. We used the same grants over the same time periods when we carried out the T tests shown in Table 3.

Looking in detail at the ACS calculations, the mean 20-day log-price return we calculated for the six ACS grant dates was equal to 32.7% while the mean over all trading days in the period from 1995 to 2002 was only 2.2%. For the mesh of 88 independent data points starting on the 18th trading date (1/19/95) we estimated that the mean and standard deviation of 20-day log-price returns were 1.4% and 11.0%, respectively, resulting in a test statistic of 6.736. Comparing to the T distribution with 87 degrees of freedom, we find that the likelihood of a test statistic this large or greater is only about 1 in 1.2 billion—roughly a factor of 250 greater than the likelihood computed by the *WSJ*.

Although the T tests for all three companies show much lower odds than the numbers estimated by the *WSJ*, we too find that it is highly unlikely that the grant dates were chosen without information about subsequent price changes. In all three cases, that hypothesis can be rejected with 99.99% confidence.[5] The main reason that the T test gives odds so much lower than those reported by the *WSJ* is the limited number of grant dates. For example, in the case of Vitesse Semiconductor, the mean 20-day log-price return on grant dates is 36% and the standard deviation (computed over the mesh starting on the first day of the period) is 23%; for nine grant dates this gives a standard error of almost 8%. Hence, the sample mean is less than five standard errors large and, after taking into account the estimated population mean (over the mesh) and the limited number of mesh data points (83), the test statistic is equal to only 4.36. It is a testament to the size of the effect that, with so few data points, the null hypothesis can be excluded with such a high degree of confidence.

6. Conclusions

The conclusions of our T test did not differ from the conclusions of the *WSJ* test for the three cases we compared. Had these cases been less extreme, however, the two tests could have led to different conclusions. Furthermore, unlike the *WSJ* test, our test is not subject to the objection that it depends on how the company made its grant date decision. The critical point is that the returns following grant dates should not be significantly higher than returns following comparable non-grant trading days. Our test makes this comparison in a statistically sound manner, yet it can be carried out easily using publicly available data.

---

[5] Rejecting a hypothesis at the 99.99% confidence level implies that the likelihood of the hypothesis being true is less than or equal to 0.01% which corresponds to odds of less than 1 in 10,000.

References

Aboody, D., Kasznik, R, 2000. CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting Economics* 29, 73–100.

Chauvin, K. W., Shenoy, C., 2001. Stock price decreases prior to executive stock option grants. *Journal of Corporate Finance* 7, 53–76.

Forelle, C., 2006. How the journal analyzed stock-option grants, *Wall Street Journal* (March 18), A5.

Forelle, C., Bandler, J., 2006. The perfect payday. *Wall Street Journal* (March 18), A1.

Heron, R. A., Lie, E., forthcoming. Does backdating explain the stock price pattern around executive stock option grants? *Journal of Financial Economics*.

Lie, E., 2005. On the timing of CEO stock option awards. *Management Science* 51, 802-812.

Pender, K., 2006. Backdating scandal has more than 100 companies in SEC's sights. *San Francisco Chronicle* (September 7).

Yermack, D., 1997. Good timing: CEO stock option awards and company news announcements. *Journal of Finance* 52, 449–476.

Year	Date	Price	20 Day Return	Annual Rank
1995	9-Mar-95	$22.50	34%	1
1996	8-Mar-96	$33.75	31%	2
1997	7-Apr-97	$21.12	30%	2
1998	8-Oct-98	$23.06	60%	1
2000	11-Jul-00	$32.88	43%	4
2002	23-Jul-02	$35.75	35%	3

Table 1: Stock option grants made to Jeffrey Rich of ACS during 1995-2002
Sources: Grant information from ACS proxy statements. Stock prices from Bloomberg. Calculations by the authors.

| **Sorted Outcome Rankings for Six Grants** | | | | | | **Number of** |
Worst	Second Worst	Third Worst	Third Best	Second Best	Best	Equivalent Outcomes
4	3	2	2	1	1	180
4	3	2	1	1	1	120
4	3	1	1	1	1	30
4	2	2	2	1	1	60
4	2	2	1	1	1	60
4	2	1	1	1	1	30
4	1	1	1	1	1	6
3	3	2	2	1	1	90
3	3	2	1	1	1	60
3	3	1	1	1	1	15
3	2	2	2	1	1	60
3	2	2	1	1	1	60
3	2	1	1	1	1	30
3	1	1	1	1	1	6
2	2	2	2	1	1	15
2	2	2	1	1	1	20
2	2	1	1	1	1	15
2	1	1	1	1	1	6
1	1	1	1	1	1	1

Total Extreme Outcomes	864
Total Possible Outcomes	256,096,265,048,064
Probability of Extreme Outcome	3.4E-12
Odds: 1 in	296,407,714,176

Table 2: The likelihood of choosing six days at random with returns ranked as high as the returns on the ACS grant dates

Company	Reported WSJ Odds (1 in X)	Computed WSJ Odds (1 in X)	Odds From T Test (1 in X)
Affiliated Computer Services	~ 300,000,000,000	296,407,714,176	726,485,163
Vitesse Semiconductor	~ 26,000,000,000	26,318,745,280	26,449
Comverse Technology	~ 6,000,000,000	6,225,819,190	26,137

Table 3: A comparison of odds computed using the *WSJ* test and our proposed T test