

---

# A Classic Case of "Data Snooping" for Classroom Discussion

Timothy Falcon Crack

*Data snooping (mistaking spurious statistical relationships for genuine ones) is an important and dangerous by-product of financial analysis. However, data snooping is a difficult concept to explain to students of financial economics because, by its very nature, it is difficult to illustrate by example (a strong statistical relationship between complex financial variables is difficult to refute). To overcome this pedagogical difficulty, I present an example of data snooping where one variable is non-financial: I show that near both new moon and full moon, stock market volatility is higher and stock market returns are lower than away from the new or full moon. The simple and off-beat nature of this example enables substantial classroom discussion.*

## INTRODUCTION

Data snooping happens when you find statistically significant but spurious relationships between empirical data. Financial analysis is particularly prone to data snooping because we use large data sets, there are many analysts looking at the same data, and we sometimes search for statistical relationships without due regard to economic justification [Lo and MacKinlay, 1990; Lo, 1994].

Data snooping is insidious when statistical relationships are sought between financial economic variables whose very complexity can mask the spurious nature of the relationship. By design, it is difficult to give classroom examples of this because it is very difficult to label a strong statistical relationship between economic variables as spurious. To overcome this problem, I present an example of data snooping where one variable is non-financial and the relationship is economically implausible: I look at the surprisingly strong statistical relationship between stock market behavior and the phases of the moon. The simplicity of the example promotes classroom understanding, and its off-beat nature makes the discussion both memorable and enjoyable.

I first set the scene by describing both the phases of the moon and the United States Naval Observatory (USNO) data used in the study. I establish the complete lack of support for lunar behavioral hypotheses in the psychology literature and then illustrate the seemingly contrary statistical relationship between phases of the moon and stock market returns and volatility. Following the analysis, I review data snooping and explain carefully why the relationships I discover are spurious. I conclude with warnings about the dangers of data snooping.

## THE SETTING: PHASES OF THE MOON

The moon moves through well defined phases.<sup>1</sup> The period of time required for the moon to go from one position relative to the sun as seen from the Earth (e.g., full moon) and back to the same position (e.g., back to full moon) is referred to as the *synodic* period. The moon's synodic period is a little over 29.5 days. However, the actual period of revolution of the moon about the Earth as seen from the stars (the *sidereal* period) is only a little over 27.3 days.<sup>2</sup>

The moon's rotation about its own axis has period of revolution equal to its sidereal period and the rotation is in the same direction. Thus, the moon always presents the same face to the Earth. Let us consider this face as a simple disk. The phases of the moon, from full to new and back to full (and the various phases within this 29.5 day cycle) may be

---

Indiana University, Bloomington, IN 47405. I thank Tom Arnold, Alex Butler, Harrison Hong and Jim Rotton.

**Table 1. Phases of the Moon (USNO 1998)**

Proportion $P$ of Disk Illuminated	Change in $P$	Phase of Moon
$P=0$	Change in $P=0$	New Moon
$0 < P < 1/2$	$P$ increasing	Waxing Crescent
$P=1/2$	$P$ increasing	First Quarter
$1/2 < P < 1$	$P$ increasing	Waxing Gibbous
$P=1$	Change in $P=0$	Full Moon
$1/2 < P < 1$	$P$ decreasing	Waning Gibbous
$P=1/2$	$P$ decreasing	Last Quarter
$0 < P < 1/2$	$P$ decreasing	Waning Crescent
$P=0$	Change in $P=0$	New Moon

Considering the moon as a disk, the proportion  $P$  of the area of the disk that is illuminated by the sun varies through the 29.5-day synodic period.  $P > 1/2$  is the gibbous phase, and  $P < 1/2$  is the crescent phase.  $P$  increases during the waxing phase, and decreases during the waning phase. The interphase period is where  $P$  is changing, that is, away from the new or full moon. Note that at new moon and full moon,  $P$  is not changing.

related directly to the proportion of the area of the disk that is illuminated by the sun and how that proportion is changing (Table 1).

The proportion of the area of the moon's disk illuminated by the sun is available from the Astronomical Applications Department of the United States Naval Observatory [USNO, 1998]. These daily data are available to two decimal places (e.g.,  $P=0.09$  on 10/19/87). My data are recorded at 12 Noon Eastern Standard Time (EST). These data are geocentric (they assume the viewer is at the physical center of the Earth), but for all practical purposes, the numbers would be the same whatever the earthly location of the observer [USNO, 1998]. These moon data are related to stock market data later in the paper.

**SPURIOUS LUNAR BEHAVIORAL HYPOTHESES**

The words "month" and "lunatic" are derived from the words "moon" and "luna," respectively. These oft-used words demonstrate how humankind has been led and misled, respectively, by the moon. The former reveals the dependence of our calendar on the motions of the moon; the latter reveals the empirically unjustified belief that lunacy is related to the motions of

the moon (an example of a "lunar behavioral hypothesis").

The typical lunar behavioral hypothesis is based on folklore that says that people behave strangely near full moon or new moon [Chapman, 1961]. There is an amazing quantity of research on all manner of lunar behavioral hypotheses. Indeed, Rotton and Kelly [1985a] and Kelly et al. [1996] review over 100 papers published in the last 30 years. These studies relate lunar movements to homicides, other criminal offences, suicides and self-harm, psychiatric disturbances, psychiatric admissions, crisis telephone calls, traffic accidents, human nutritional intake and many other human phenomena. Without exception, these studies find no evidence whatsoever to support any link between the motion of the moon (however defined) and any form of human behavior. The bottom line is that there is zero empirical support for any lunar behavioral hypothesis.

With no empirical support, some researchers have turned instead to looking at peoples' *beliefs* about lunar behavioral hypotheses. Many people still believe that the moon influences behavior [Rotton and Kelly, 1985b; Wilson and Tobacyk, 1990]. Such beliefs are dangerous because they may adversely affect the way mental health workers respond to patients [Vance,

**Table 2. Results for the CRSP EWRETD (equally weighted) Index**

"	$N_{NF}$	$N_I$	$253\hat{\mu}_{NF}$	$253\hat{\mu}_I$	Z-stat ( <i>p</i> val.)		$F_{NF}\sqrt{253}$	$F_I\sqrt{253}$	Z-stat ( <i>p</i> val.)		Joint $P^2$ -stat ( <i>p</i> val.)
0.20	2948	2108	0.2121 (0.0527)	0.2812 (0.0453)	-1.389	(0.083)	0.1062 (0.0141)	0.0920 (0.0058)	1.461	(0.072)	2.837 (0.242)
0.25	3334	1722	0.2115 (0.0508)	0.2978 (0.0462)	-1.734	(0.042)	0.1045 (0.0130)	0.0923 (0.0061)	1.428	(0.077)	3.647 (0.162)
0.30	3703	1353	0.2076 (0.0491)	0.3319 (0.0519)	-2.219	(0.013)	0.1037 (0.0120)	0.0911 (0.0068)	1.844	(0.033)	5.339 (0.069)
0.35	4056	1000	0.2215 (0.0478)	0.3196 (0.0540)	-1.690	(0.046)	0.1024 (0.0115)	0.0926 (0.0075)	1.645	(0.050)	3.326 (0.190)
0.40	4390	666	0.2280 (0.0453)	0.3259 (0.0616)	-1.644	(0.050)	0.1018 (0.0116)	0.0916 (0.0057)	1.025	(0.153)	2.834 (0.243)

These index data are drawn from the combined NYSE/AMEX and NASDAQ data. The continuously compounded daily index returns data span the 20 year period January 1977 to December 1996. The subscript "NF" denotes the New/Full Period ( $P < "$  or  $P \geq 1-$ "); the subscript "I" denotes the Interphase period (" $\#P < 1-$ "). The means  $\mu_i$  and standard deviations  $F_i$  have been annualized using a factor of 253 (standard errors are in parentheses). The sample sizes are denoted as  $N$ . Three test statistics are reported: one for the test of  $\mu_{NF} = \mu_I$ , one for  $F_{NF} = F_I$ , and one for the joint test of both hypotheses. The first two are standard normal (technically, I should square them and report them as chi-squared with one d.f.), the last is chi-square with two d.f. Each test is one-sided (the alternative is that the NF period has higher volatility and lower mean than the I period). The statistics are calculated using GMM with autocorrelation and heteroskedasticity robust standard errors.

1995] or the way police officers react to criminal activity [Rotton et al., 1986]. If *financial analysts* hold such beliefs and act on them, they risk poor asset allocation, poor risk-adjusted portfolio performance, and ridicule. Of course, financial analysts face the same consequences when basing their decisions on any spurious analysis - lunar or otherwise.

### THE MOON AND THE STOCK MARKET

The null hypothesis is common sense: phases of the moon are irrelevant to the stock market. The alternative lunar behavioral hypothesis is consistent with folklore [Chapman, 1961 and Kelly et al., 1996]: the period of time near the new moon or the full moon is "bad" relative to the interphase period (the period away from new and full moon). In particular, I hypothesize that investors like returns and dislike volatility, so my lunar behavioral hypothesis is simply that near new moon or full moon, volatility will be

higher and returns will be lower than during the interphase period.

I use 20 years of CRSP daily data: 01/01/77-12/31/96. I look at the CRSP indices EWRETD and VWRETD (equally-weighted with dividends and value-weighted with dividends).<sup>3</sup> The two indices are drawn from the combined NYSE/AMEX and NASDAQ data. I use continuously compounded returns (the log of the "price relative").

I isolate periods near the full moon and new moon using the parameter  $P$  (the proportion of the face of the moon illuminated by the sun). When  $P$  is near zero or near one, we are close to the new and full moons, respectively. I use a parameter  $\theta, 0 < \theta < 1/2$ , to separate the data into a "New/Full Period" ( $P < \theta$  or  $P \geq 1-\theta$ ), and an "Interphase Period" (" $\#P < 1-\theta$ ") using dummy variables. The hypotheses are tested for various  $\theta$  values because it is not clear *a priori* which value gives the most powerful tests. All test statistics are calculated using GMM with Newey-West standard errors. The tests are therefore robust to both

**Table 3. Results for the CRSP VWRETD (value weighted) Index**

"	$N_{NF}$	$N_I$	$253\hat{\mu}_{NF}$	$253\hat{\mu}_I$	Z-stat ( $p$ val.)	$F_{NF}\sqrt{253}$	$F_I\sqrt{253}$	Z-stat ( $p$ val.)	Joint $P_2^2$ -stat ( $p$ val.)
0.20	2948	2108	0.1018 (0.0454)	0.1812 (0.0468)	-1.148 (0.126)	0.1392 (0.0189)	0.1216 (0.0074)	1.294 (0.098)	2.497 (0.287)
0.25	3334	1722	0.1030 (0.0424)	0.1977 (0.0500)	-1.343 (0.090)	0.1373 (0.0175)	0.1213 (0.0077)	1.315 (0.094)	2.851 (0.240)
0.30	3703	1353	0.1019 (0.0386)	0.2265 (0.0539)	-1.794 (0.036)	0.1350 (0.0161)	0.1238 (0.0090)	1.129 (0.130)	3.592 (0.166)
0.35	4056	1000	0.1194 (0.0394)	0.1995 (0.0568)	-1.092 (0.152)	0.1338 (0.0155)	0.1253 (0.0093)	0.887 (0.188)	1.254 (0.534)
0.40	4390	666	0.1253 (0.0354)	0.2013 (0.0705)	-0.898 (0.185)	0.1329 (0.0150)	0.1289 (0.0098)	0.377 (0.353)	0.837 (0.658)

These index data are drawn from the combined NYSE/AMEX and NASDAQ data. The continuously compounded daily index returns data span the 20 year period January 1977 to December 1996. The subscript "NF" denotes the New/Full Period ( $P < "$  or  $P \leq 1-$ "); the subscript "I" denotes the Interphase period (" $\# P < 1-$ "). The means  $\mu_i$  and standard deviations  $F_i$  have been annualized using a factor of 253 (standard errors are in parentheses). The sample sizes are denoted as  $N_i$ . Three test statistics are reported: one for the test of  $\mu_{NF} = \mu_I$ , one for  $F_{NF} = F_I$ , and one for the joint test of both hypotheses. The first two are standard normal (technically, I should square them and report them as chi-squared with one d.f.), the last is chi-square with two d.f. Each test is one-sided (the alternative is that the NF period has higher volatility and lower mean than the I period). The statistics are calculated using GMM with autocorrelation and heteroskedasticity robust standard errors.

autocorrelation and heteroskedasticity in the moments. The results of the hypothesis tests for the two indices are reported in Tables 2 and 3, respectively. Looking across different " values in Tables 2 and 3, we see that *without exception*, the index mean return is lower and the index volatility is higher near new and full moons than during the interphase period. The results for the EWRETD index are, in every case, statistically stronger than the results for the VWRETD index. Small capitalization stocks have a lesser influence on the VWRETD index than on the EWRETD index, so this suggests some sort of size effect, where small stock returns are less likely to satisfy the null hypothesis.

Cutting the  $P$  values at " $=0.30$  gives the most significant results. In this case, the hypothesis that the mean returns are equal across phases of the moon is rejected with  $p$ -values of 1.6% for the EWRETD and 3.6% for the VWRETD. The hypothesis that volatility is the same across phases of the moon is rejected for the EWRETD with a  $p$ -value of 3.3%, but is only

marginally significant for VWRETD because the  $p$ -value is 13%. The tests for differing volatility produce weaker results than those for differing mean. The joint test of mean and volatility is significant for EWRETD ( $p$ -value 6.9%) and not really significant for VWRETD ( $p$ -value 16.6%).

On the face of it, it appears that the stock market lunar behavioral hypotheses is strongly supported for means and weakly supported for variances. There also appears to be a size effect, where the null is less likely to be true for smaller capitalization stocks.<sup>4</sup>

### IT IS JUST DATA SNOOPING

It is well known that if you look hard enough within a large data set, you find spurious relationships. For example, I am told that in The Netherlands, changes in the human birth rate are statistically significantly related to changes in the stork population (yes, the big birds that deliver babies in fables).

---

In financial economics, looking at data that others have looked at can lead to severe inference problems [Lo and MacKinlay, 1990; Lo, 1994]. For example, suppose a researcher runs 10,000 regressions relating abnormal CAPM performance to every economic variable imaginable. The researcher may find that of the 10,000 regressions, abnormal performance is most closely related to some Factor X (unjustified by economic reasoning). If another researcher forms decile portfolios sorted on Factor X, this effectively sorts the abnormal performance t-statistics into a pseudo-ordering determined via this spurious Factor X [Lo and MacKinlay, 1990]. Using the t-statistic for the extreme Factor X portfolio is akin to running many regressions yourself and taking the one with the largest t-statistic. This biases statistical inferences by increasing the size of tests (it also increases the power).

In the case of the moon data and the lunar behavioral hypotheses, something very similar is happening. Many lunar behavioral hypotheses coexist in the psychology literature (in the absence of any empirical data to confirm any one in particular). Some hypotheses emphasize the full moon, others the new moon, others the waxing or waning. Some emphasize the first quarter or last quarter, the distance of the moon from the Earth, the synodic period or sidereal period (there are three other periods I do not mention), while others look at combinations of these [Rotton and Kelly, 1985a]. The muddled psychology literature leaves the door open for a researcher to explore many possible lunar hypotheses until one (the "New/Full Period" versus "Interphase Period" hypothesis) gives strong statistical results. This is data snooping - the results are spurious.

How do we know that the results really are spurious? There are several ways to tell. First, any economic justification of the results requires as a prerequisite that there is evidence of a link between the motion of the moon and human behavior. However, thirty years of psychology literature on lunar behavioral hypotheses finds no evidence whatsoever of such a link (see my earlier section). Second, I have confessed that in the absence of any economic arguments, I had to snoop through different possible relationships before I found one that was significant and then I had to choose  $\beta = 0.30$  in particular to get

very significant results (such snooping completely destroys the meaning of p-values). Third, I am using twenty years of data because the results are not as significant for shorter time series - an example of Lindley's Paradox.<sup>5</sup> Finally, the 1987 stock market crash happened just after a new moon and removing October 1987 from the data also removes the statistical significance of volatility differences (without changing the significance of mean return differences). Of course, true believers might argue that the Crash of 1987 should be left in the data precisely because it supports the alternative hypothesis.

In the absence of any economic justification, and given my data-snooping confessions, these results cannot be interpreted as anything other than an elegant and off-beat example of data snooping. Unfortunately, most data snooping in financial economics is much more insidious - because plausible economic variables are typically involved. Whatever the data, accepting a statistical relationship in the absence of sound economic justification exposes you to potential data snooping biases with potentially dangerous consequences. For example, anyone who treats the results of this study as economically meaningful might conclude that we must reject the CAPM because it fails to incorporate a moon factor. Such a deduction is sheer lunacy!

## CONCLUSION

Data snooping is dangerous because spurious statistical relationships can be mistaken for meaningful economic ones. However, by its very nature, data snooping is difficult to explain in class by example because strong statistical relationships between financial economic variables are not easily labeled as spurious. I overcome this difficulty by examining a non-financial variable (the phases of the moon) and relating it to both stock market returns and volatility. Although returns and volatility in the stock market are both strongly statistically related to phases of the moon, and although the results are consistent with folklore, these results are shown to be spurious and generated by data snooping.

The combination of a very strong statistical relationship with the complete absence of economic plausibility hammers home the point that statistical relationships not well-founded in economic theory are extremely dangerous.

---

## ENDNOTES

<sup>1</sup>This section draws heavily on Considine [1976].

<sup>2</sup>The moon's synodic period (e.g., full moon to full moon) is longer than the sidereal period because the Earth moves about the sun in the same direction as the moon. By the time the moon has gone 360 degrees around the Earth, the position of the Earth relative to the sun has changed. It takes the moon an extra 2.2 days to get back to the same position relative to the sun as seen from the Earth. A simple formula relates the moon's sidereal period  $S$ , its synodic period  $N$ , and the Earth's sidereal period  $E$  (approximately 365.25 days):  $1/S=1/N+1/E$ .

<sup>3</sup>Rotton and Kelly [1985a, p302] mention an unpublished paper by Rotton and Rosenberg [1994] who look at five years of Dow Jones Industrial Average (DJIA) data. They find that the *level* of the DJIA is related to the lunar cycle. However, the result disappears when the data are first differenced and "pre-whitened."

<sup>4</sup>These results should be independent of any non-lunar calendar anomalies [e.g., Cross, 1973 or Roll, 1983] because the moon's synodic period of 29.5 days is not an integer multiple of 7 (days in a week) and cannot be multiplied by an integer to get 365.25 (days in a year).

<sup>5</sup>Lindley's Paradox says (essentially) that if the null hypothesis is incorrect by even a very slight margin, then as long as you use a large enough data set, you can get very significant results [Lindley, 1957].

## REFERENCES

- Chapman, Loren J. "A Search for Lunacy," *Journal of Nervous and Mental Disease*, 132 (1961), 171-174.
- Considine, Douglas M. (editor). *Van Nostrand's Scientific Encyclopedia* (New York, NY, 1976), Van Nostrand Reinhold Company.
- Cross, Frank. "The Behavior of Stock Prices on Fridays and Mondays," *Financial Analysts Journal*, 29 (1973), 67-69.
- Kelly, I.W., James Rotton, and Roger Culver. "The Moon was Full and Nothing Happened: A Review of Studies on the Moon and Human Behavior and Human Belief," in *The Outer Edge: Classic Investigations of the Paranormal*, edited by Joe Nickell, Barry Karr and Tom Genoni, (Amherst, NY, 1996), Committee for the Scientific Investigation of Claims of the Paranormal: 16-34.
- Lindley, D.V. "A Statistical Paradox," *Biometrika*, 44 (1957), 187-192.
- Lo, Andrew W. "Data-Snooping Biases in Financial Analysis," In *Blending Quantitative and Traditional Equity Analysis*, (Charlottesville, VA, 1994), Association for Investment Management and Research: 59-66.
- Lo, Andrew W., and A. Craig MacKinlay. "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3 (1990), 431-468.
- Roll, Richard. "Vas Ist Das: The Turn of the Year Effect and the Return Premia of Small Firms," *Journal of Portfolio Management*, 9 (1983), 18-28.
- Rotton, James, and I.W. Kelly. "Much Ado About the Full Moon: A Meta-Analysis of Lunar-Lunacy Research," *Psychological Bulletin*, 97 (1985a), 286-306.
- Rotton, James, and I.W. Kelly. "A Scale for Assessing Belief in Lunar Effects: Reliability and Concurrent Validity," *Psychological Reports*, 57 (1985b), 239-245.
- Rotton, James, I.W. Kelly, and Patricia Elortegui. "Assessing Belief in Lunar Effects: Known-Groups Validation," *Psychological Reports*, 59 (1986), 171-174.
- Rotton, J., and M. Rosenberg. "Lunar Cycles and the Stock Market: Time-Series Analysis for Environmental Psychologists," Unpublished manuscript, (Florida International University, North Miami, Florida, 1984).
- Vance, David E. "Belief in Lunar Effects on Human Behavior," *Psychological Reports*, 76 (1995), 32-34.
- USNO. Information culled from the website of the Astronomical Applications Department of the United States Naval Observatory (a department of the Navy). 1998. Web Reference: <http://aa.usno.navy.mil/AA/>