Beta Estimation with Stock Return Outliers:

The Case of U.S. Pharmaceutical Companies

Alexandra Theodossiou*

Panayiotis Theodossiou†

Uzi Yaari‡

Revised May 20, 2009

Abstract

Efficient estimation of the equity cost of public corporations is an essential component of calculating the required rate of return of real investment projects, and therefore the basis for a rational investment policy. The accepted methodology relies on the CAPM model to define the return risk premium, and the OLS method to estimate the beta risk coefficient required for calculating the premium. This study challenges the use of the OLS method for this task by demonstrating its vulnerability to the impact of stock return outliers caused by large, unpredictable, company-specific events. That impact is verified on a sample of U.S. pharmaceutical companies by comparing the OLS estimation performance with that of our proposed method based on Huber's Robust M (HRM) estimator, a related statistical method that follows a mixed return model identifying regular and outlier return components. Using the HRM-estimated beta as a benchmark, we demonstrate that (1) outliers can substantially bias the OLS beta, (2) the bias is negatively correlated with company size, and (3) the size of the bias is often moderated but not eliminated by extending the estimation period. The latter finding suggests that a robust method like HRM is preferable where estimators ought to represent the behavior of the majority of historical data despite the presence of outliers. The risk of trusting the OLS beta is especially high when estimation must rely on a small sample.

JEL: G12

Keywords: OLS estimation; Robust M estimation; Stock beta; Pharmaceutical industry

Preliminary and Incomplete Version: Please do not quote without permission

---

* Dr. Alexandra K. Theodossiou; Email: atheodoss@gmail.com

† Dr. Panayiotis Theodossiou is Professor of Finance at Rutgers, The State University of NJ, 227 Penn Street, Camden, NJ 08102-1519; Email:theodoss@camden.rutgers.edu.

‡ Dr. Uzi Yaari is Professor of Finance at Rutgers, The State University of NJ, 227 Penn Street, Camden, NJ 08102-1519; Email:yaari@camden.rutgers.edu.

I. Introduction

The U.S. pharmaceutical industry is characterized by a symbiotic relationship between a few old giants and numerous newer and smaller companies. The former specialize in manufacturing and marketing of established products and occasional acquisition of rights over promising new products. New products are researched and developed by risk-taking small companies in the expectation of an eventual acquisition by one of the giants. News announcements concerning pharmaceutical products, such as the outcome of a clinical trial, a regulatory decision, or a class action lawsuit, can have a dramatic impact on the stock price of a company with consequent return outliers. Although few and far between, outliers have the power to significantly alter a company's estimated beta.

A documented drawback of the OLS method in estimating the CAPM parameters of a single company lies in the sensitivity of this method to deviations from normal distribution, in this case due to outlier returns. The impact of outlier returns in this context is investigated by Martin and Simin (2003), Gray et al. (2005), and Genton and Ronchetti (2008).[1] Evidence shows that in an extreme case, even a single outlier may significantly affect the CAPM regression estimates of a stock's alpha, beta, and variance. A misstated beta would lead to a misstated cost of capital, risking erroneous capital budgeting decisions. In view of evidence by Brunner et al. (1998) and Graham and Harvey (2001) that over three-quarters of the U.S. companies rely on the CAPM to estimate their cost of capital, a disregard of this risk may entail a large-scale misallocation of resources and attached social cost. Across industries, the extent of the damage is likely to rise with the severity of return outliers.

One approach of mitigating the impact of outliers on estimated OLS parameters calls for the removal of those observations from the sample. Applying this procedure to a sample of Australian firms, Gray et al. (2005) obtain stock return models that outperform the unadjusted OLS beta, the unity beta

[1] Martin and Simin (2003) and Genton and Ronchetti (2008) demonstrate the impact of outliers on the estimated parameters of the single-factor model. Gray et al. (2005) provide examples of such an impact on stock return models using Australian utility companies.

model, and Blume's (1975) adjusted-beta model. Under another approach, outlying observations are preserved while explanatory variables are individually Winsorized (trimmed): highest and lowest percentages are raised or brought down to predetermined values. Using the Least-Trimmed Squares (LTS) methodology, Knez and Ready (1997) find that the Fama and French (1992) size-based risk premium disappears.

The elimination of outlier observations and the trimming of outlier explanatory variables have both been criticized by statisticians. Trimming alters the stochastic relationship among explanatory variables when the outlier component is correlated with other explanatory variables. Omitted observations reduce the size and statistical power of the estimated model. Alternative solutions that preserve outliers may yield better if not perfect measures of systematic risk.

This paper proposes a flexible theoretical model of stock return with an outlier component, which follows Huber's (1964, 1981) Robust M (HRM) estimator. The mathematical relationship of our model to the OLS provides a framework for measuring and testing the impact of return outliers on beta.[2] Examined are also the related questions of whether this impact is related to company size, and to what extent this impact can be mitigated or even eliminated by extending the estimation period.

The paper proceeds as follows. Section II describes the economic importance of the pharmaceutical industry in terms of its size, growth, and profitability. Section 3 details our sample and preliminary empirical results. Section 4 identifies outlier-provoking company-specific events of four companies and the magnitude of market reaction to those events. Section 5 examines the relationship between sample size and the OLS bias. Section 6 offers a summary and concluding remarks.

[2] The mathematical relationship between the OLS and the HRM return-generating estimators are analyzed in P. Theodossiou and A. K. Theodossiou (2009), Impact of Outliers on Stock Return Models: Implications for Event Studies and the Pricing of Risk, *Working Paper*. Those relations allow the use of the HRM methodology even when outlier returns should be measured at full length.

II. The Pharmaceutical Industry

Between 1960 and 2006 overall U.S. national health expenditures grew from $27.5 billion to $2.1 trillion or from 5.2 percent to 16.0 percent of GDP. [3] By 2017 those figures are projected to reach $4.3 trillion representing 19.5 percent of GDP. Over the same period, the use of prescription drugs alone grew from $2.7 billion to $216.7 billion, and they are projected to reach $515.7 billion in 2017. While the projected average annual growth rate of the national health expenditures is 6.7 percent, prescription drugs are expected to grow at the annual rate of 8.2 percent (Keehan et al., 2008). As a percentage of GDP, prescription drugs are expected to increase from 1.64 percent in 2006 to 2.35 percent in 2017. These statistics leave no doubt that the pharmaceutical industry is an important component of the US economy and increasingly so.

Fortune Magazine consistently ranks the pharmaceutical industry as one of the most profitable in the U.S. based on its profit-to-revenue ratio. This industry was ranked the most profitable between 1995 and 2002, the third most profitable between 2003 and 2004, the fifth most profitable in 2005, and the second most profitable in 2006. In 2007 it was ranked third with a profit-to-revenue ratio of 15.8 percent, a figure surpassed only by Network and other Communications, Mining, and Crude Oil Production. In the same year the Medical Products and Equipment Industry was ranked fourth with a profit-to-revenue ratio of 15.2 percent. In 2008 the pharmaceutical industry again ranked third and the Medical Products and Equipment Industry fourth with profit to revenues ratios of 19.3 and 16.3 percent, respectively.

Under the current political climate, a new government policy under consideration would rely on enhanced regulation and competition to limit the secular rise in health-care profits and costs to the growth rate of GDP.[4] A key ingredient of such a regulation would be an objective estimate of companies' cost of capital similar to the treatment of public utilities.

[3] Source: www.cms.hhs.gov, Centers for Medicare and Medicaid Services.
[4] Source: www.kff.org, Comparing Projected Growth in Health Care Expenditures and the Economy, May 2006

III. Sampling and Preliminary Results

Our sample consists of 76 pharmaceutical companies reported on the CRSP database for which complete data are available for ten consecutive years between January 1980 and August 2008. Those companies engage in research and production of medicinal chemicals and botanical products, pharmaceutical preparations, in-vitro and in-vivo diagnostic substances, and biological products (SIC codes ranging from 2830 to 2836 in both CRSP and COMPUSTAT). Thirty-eight companies are excluded from the sample for having less than ten years of continuous monthly data.

Table 1 reports company statistics on market capitalization (i.e., common stock value), mean and standard deviation of monthly stock returns, and stock beta estimates based on the OLS and Huber's Robust M (HRM) methods. Companies are sorted by median market capitalization based on monthly figures over the five-year sample window, from September 2002 to August 2007. Basic statistics at the bottom of the table refer to the sample distribution of key measures reported in the table.

Results reveal that the mean market capitalization of the individual companies sampled ranges from $6 million to $196 billion, and their combined capitalization (hereafter industry capitalization) is $706 billion. The combined capitalization of the largest five firms (Pfizer, Merck, Amgen, Abbot, and Lilly) is $502.6 billion, constituting 71.2% (= 502.6 / 706) of the industry capitalization. The combined capitalization of the largest 10 and 20 companies is $637.6 and $689.3 billion, making up 90.3% and 97.6% of the industry, respectively. In contrast, the smallest 5, 10, and 20 companies have a respective combined capitalization of $17.8, $39.8, and $94.1 million, constituting 0.009%, 0.036%, and 0.128% of the industry. Eighty percent of the companies have a market capitalization ranging from $39 million to $16.29 billion, and fifty percent of the companies from $76 million to $1.869 billion. The median company capitalization is $229.9 million.

The second and third columns in Table 1 report the effect of the individual company on the cumulative share of industry capitalization in the form of decrement (second column) and increment (third column). As indicated by the summary statistics at the bottom of the table, 50%, 75%, and 90% of the smallest companies have a combined market capitalization of 0.51% (3.6 BB$), 2.88% (20.3 BB$), and 16.4% (116 BB$), respectively, in reference to the industry total of 706 BB$.

The distribution of market capitalization is further analyzed using a Lorenz curve and its Gini coefficient. Displays in Figure 1, the Lorenz curve shows the cumulative percentage of capitalization contributed by any cumulative share of the smaller companies. It measures company size inequality in reference to the special case of a uniform distribution represented by a 45-degree diagonal line. The Gini coefficient reduces to a single ratio the information provided by the Lorenz curve. The area bounded between the Lorenz curve and the diagonal line of perfect equality is divided by the entire triangular area under the diagonal line. Perfect equality would be indicated by a ratio of zero, and perfect inequality by a ratio of one. Our Gini coefficient of 0.891 is consistent with the reality of an industry dominated by the performance of a small number of relatively large companies.

The fourth and fifth columns in Table 1 report the individual company average and standard deviation of monthly returns over the five-year sample period. The minimum, median, and maximum of the average monthly returns are -2.7% (Genta), 1.9%, and 7.2% (Advanced Magnetic), respectively. The average monthly return ranges between −1% and 5.5% for eighty percent of the companies, and between 0.5% and 2.8% for fifty percent of the companies. Sixty seven companies exhibit a positive average return, and nine companies a negative average return. When compounded, a monthly return of 1.94% is equivalent to an annual return of 25.97%, figures far higher than the average monthly return for the market as a whole, 0.97%, or it annual equivalent, 12.26%.

The minimum, median, and maximum standard deviation of monthly returns are, respectively, 5.1% (Abbot), 17.3% (Isis), and 51.3% (Genaera). The contemporaneous figure for the stock market at large is 4.02% while for eighty percent of the companies the standard deviation is between 6.8% and 26.4%, and for fifty percent of the companies it is between 10.3% and 21.7%. Table 1 further reveals that large-capitalization companies exhibit lower standard deviations of returns than smaller companies.

The last three columns in Table 1 report estimates of each company’s stock beta based on the OLS and HRM methods, and the absolute difference between the two. The respective minimum, median, and maximum absolute differences between the two beta estimates are 0.001, 0.229, and 2.103. Roughly, twenty-five percent of the companies show a difference larger than 0.509 and ten percent show a difference larger than 0.724, where small companies average a larger difference than large companies and the average difference in large companies is absolutely small. To further examine the relationship between company size and OLS bias, we run the regression

$$\left|\beta_{OLS,i} - \beta_{k,i}\right| = \underset{(6.16)}{0.7292} - \underset{(-3.4)}{0.061}\ln\left(MVE_i\right), \quad R^2 = 0.1349\,,$$

where $\left|\beta_{OLS,i} - \beta_{k,i}\right|$ is the absolute difference between the OLS and HRM beta estimators for company i, ln(MVE) is the natural logarithm of the company's market capitalization used as proxy for size, and i =1, 2,…,76 identifies the company. The estimated regression slope is negative and statistically significant at the 1% percent level, confirming the hypothesis that the estimators' absolute difference is negatively related to company size. This finding is further supported in the next section where we demonstrate the greater price sensitivity of smaller companies to announcement effects.

IV. OLS vs. HRM: Four Company Case Studies

The following examples provide preliminary evidence of the potential benefit from using the HRM estimator for calculating stock betas, especially in the case of growth stocks where company-specific events are more likely to induce large price reactions. The CRSP and Factiva databases are the sources of data on the major news items behind a stock's outlying monthly returns during 9/2002–8/2007. In the tables below, the Date identifies the year and month (yyyy/mm) of a winsorized outlying return, MV is the end-of-month aggregate common stocks value in millions of dollars, Price is the monthly closing share price in dollars, Shares represent the end-of-month number of outstanding common shares in millions, Vol is the monthly trade volume in millions of shares, and Ri and Rm, respectively, are the company stock and market-wide monthly returns in percentage. During the sampling period, market-wide monthly returns varied from -10% to 8%.

Genaera

Founded in 1987 and headquartered in Plymouth Meeting, Pennsylvania, Genaera Corporation is a biopharmaceutical company that engages in the research and development of pharmaceuticals with a focus on obesity, anti-angiogenesis, and respiratory diseases. During the sampling period, the company's market capitalization ranges from $11 to $233 million with a monthly median value of $97 million.

Table 2 reveals that the company's monthly return volatility (column 5) is considerably higher than that of the overall market (column 6) with the former sinking as deep as −45% and soaring as high as +339%. We note that these figures pertain to the company's total price volatility combining systematic and unsystematic risk.

Figure 2A displays the two regression lines side-by-side. Compared with the moderate beta of 0.9944 based on the HRM estimator, the company's OLS beta is quite high at 3.123. We observe that the steeper OLS regression slope in Figure 2A is driven by a few very large positive returns during 5/2005–7/2007 (Table 2, column 5). The abnormal returns are clustered in the periods 10/2002–8/2003 and 1/2007–4/2007 corresponding to the following company-specific events and subsequent announcements.

- October 2002: A decrease in EPS during the first six months of the year compared to the same period in the previous year. A price decrease of 41.5 percent.
- December 2002: Regulatory approval by the Irish Medicines Board to begin a Phase-II clinical trial of Lomucin, an oral treatment for cystic fibrosis. A price increase of 36.2 percent.
- January 2003: Cancellation of new enrollment in the Phase-I/II clinical trial of the drug Squalamine in its use for treating non-small-cell lung cancer. A price decrease of 39.1 percent.

- March 2003: Nasdaq extends by 180 days the grace period for compliance with the minimum requirement of a $1 bid price, a condition for the company's continued listing on the Exchange. A price increase of 90.6 percent.
- April 2003: Correction of the previous decision. A price decrease of 32 percent.
- May 2003: Phase-I/II clinical trial of the drug Squalamine in its use for treating macular degeneration shows improved vision of some patients and stabilizes vision of all others. Later in the month: Based on preclinical testing, two other candidate drugs may be effective in treating asthma. A price increase of 33.9 percent. This event is largely responsible for an upward shift in the OLS regression line and an increase in the line's slope (beta) to 3.123.
- July 2003: Promising results of the drug Squalamine in its phase-I/II clinical trial of treating non-small-cell lung cancer. Later that month, the European Patent Office grants the company a patent over the use of Squalamine to inhibit the growth of endothelial cells. A price increase of 61 percent.
- August 2003: Evidence that Squalamine improves vision in age-related macular degeneration. Later that month, two related US patents are issued. A price increase of 63.6 percent.
- January 2004: A successful private placement of 4,950,500 common shares at $4.04 per share. The deal includes 990,100 warrants to purchase 990,100 additional shares. A price increase of 43.7 percent.
- August 2005: Executive VP John L. Armstrong is promoted to President and Chief Operating Officer. A price increase of 52.1 percent.
- June 2006: Patent received on gene variants of the Interleukin-9 receptor. The company is given 180 days to regain compliance with Nasdaq's minimum bid requirements. A price decrease of 44.7 percent.
- January 2007: Stopped development of Evizon, a drug aimed at treating wet, age-related macular degeneration, in the middle of its testing on patients. The company lays off 30 percent of its work

force. A price decrease of 34.24 percent. Despite the absence of major news, this change is reversed in February 2007. A price increase of 52.06 percent.

- April 2007: A new CFO is hired to concentrate investment in the development of Trodusquemine, a drug for the treatment of obesity. Later in the month, a one-for-six reverse split of common stock. A price increase of 35 percent.

Inspection of Figure 2A reveals that large positive or negative returns triggered by company-specific events are atypical of company returns and come from a different distribution. Dropping such observations from the sample would lead to a loss of information, while ignoring their impact would result in a volatile, inefficient estimate of the regression slope thereby masking the true relationship between company and market returns over the relevant long-term horizon. Conceptually, a stock beta is proportional to the correlation between the stock return and that of the market. Since the HRM estimator is largely immune from the effect of extreme returns due to idiosyncratic company events, we propose that it better represents the long-term relationship between the return of a stock and that of its market.

Amylin

Founded in 1987 and based in San Diego, Amylin Pharmaceuticals is a biopharmaceutical corporation that engages in research, development, and commercialization of medicines for diabetes, obesity, and other diseases. The company's market capitalization over the sampling period ranges from \$1.32 to \$6.48 billion with a median value of \$2.29 billion. Monthly return varies considerably between −18% and 76%. The OLS beta is negative at −0.3345 compared with a positive HRM beta of 0.355. Figure 2b reveals that the regression sign reversal under OLS is caused by large positive returns during a declining market in September 2002 and August 2005 (Table 2, columns 5-6).

- September 2002: The company stock price temporarily increases by 37.8% on news of its collaboration with Eli Lily to produce and market a new generation therapy for type-2 diabetes. A price decrease of 10 percent.

- June 2005: Amylin and Eli Lily announce the launch of Byetta, an injection drug treatment for type-2 diabetes. A price increase of 31 percent.
- August 2005: the company announces promising results from an ongoing Phase-II study of Byetta. A price increase of 75.5 percent.

Figure 2B reveals that the large positive stock returns during negative market returns in September 2002 and August 2005 reversed the sign of the OLS-estimated beta. Of the two positive observations, that of a smaller return has a greater impact because of the underlying negative market return. As in the previous example, the HRM regression provides a better representation of the long-run return tendency.

Poniard

Originally known as NeoRx, Poniard Pharmaceuticals (Figure 2C) was founded in 1984 and changed its name in 2006. This biotechnology company is involved in research, development, and commercialization of cancer therapy products in the U.S. The company's market capitalization over the sampling period ranges from \$10.4 to \$282.1 million with a median value of \$74 million.

Monthly returns vary dramatically between −44% and 204%. Consistently, the OLS beta estimator is high at 5.2208, and so is the lower HRM beta at 3.5960. The steeper OLS regression slope is driven by two large positive returns in May and September 2003.

- October 2002: Spreading news that Nasdaq is looking to help companies like NeoRx that are in danger of delisting. A price increase of 66.9 percent.
- March 2003: Appointment of Jack Bowman as Executive Chairman and transfer of the company's listing from Nasdaq's national market to its small-cap market. A price increase of 55.1 percent.
- April 2003: The FDA lifts its hold on the clinical use of its Skeletal Targeted Radiotherapy for the treatment of multiple myeloma, a cancer of the bone marrow. A price increase of 204 percent.

- September 2003: Spreading analysts' news that the Black Box[5] generated a "buy" signal for the company's stock. The company undertakes conference calls to promote its Skeletal Targeted Radiotherapy product. A price increase of 120.3 percent.
- February 2006: An infusion of $65 million of private capital, a move which is expected to prevent insolvency and prolong the company's life for five years at its present spending rate. A price increase of 67.8 percent.
- November 2006: Encouraging survival results in a phase-II trial of Picoplatin, a drug for the treatment of small-cell lung cancer. Later that month, the beginning of a phase-III trial for the SinuNase drug. The company regains its compliance with Nasdaq's minimum bid requirement. A price increase of 55.1 percent.

Merck

One of the world's largest pharmaceutical companies, Merck & Co., Inc., was founded in 1891 and incorporated under its present name in 1927. The company is engaged in the research, development, manufacturing and marketing of human and animal medicines and vaccines. Research focus is on therapies for Alzheimer's, atherosclerosis, cancer, diabetes, cardiovascular disease and treatments for pain and sleep disorders. Market capitalization over the sampling period ranges from $59.74 to $135.8 billion with a median value of $96.892 billion. Monthly returns vary between −26% and 19%. The OLS beta is 0.7255 compared with an HRM beta of 0.7573 (Figure 2D).

- October 2003: Reported earnings per share for the third quarter are $0.83 compared with $0.85 forecasted by Wall Street. In the same month the company announced the elimination of 4,400 jobs from its workforce. A price decrease of 12.6 percent.

[5] The Black Box is a computer investing program consisting of a series of formulas that process data entered by investors to formulate optimal trading strategies. As suggested by its name, users need not understand the algorithm generating the results.

- September 2004: A worldwide voluntary withdrawal of the company's best-selling drug, Vioxx, following a study linking its use with a higher risk of heart failure and stroke. A price decrease of 25.8 percent.
- November 2004: Expanding troubles with Vioxx, related litigations, and an FDA request for additional information about the efficacy and safety of the company's new drug, Arcoxia. A price decrease of 10.5 percent.

Contrary to the previous examples, in the case of Merck the OLS and HRM beta estimators produce similar results. The three winsorized observations have no significant impact on the estimated beta. This result is consistent with the observation that Merck's stock return is stable, presumably a consequence of the greater internal diversification supported by the company's larger size.

V. OLS and HRM vis. Sample Size

Real-life constraints often limit the period of measurement and related sample size. The samples we used so far to estimate individual stock betas consist of sixty consecutive monthly observations of a stock returns paired with those of the underlying market. In a modest sample, the presence of a few outlying observations may have a significant impact on the OLS-estimated CAPM parameters. The impact on beta would be mitigated in a larger sample provided outliers are randomly distributed in both tails of the return distribution.

To compare the power of the OLS and HRM methods to cope with outliers, we use side-by-side samples of different sizes where results based on the larger sample offer a benchmark for both methods. For each company under each sample size, we estimate beta using both methods and calculate the difference $\beta_{OLS,i} - \beta_{k,i}$, hereafter referred to as the OLS bias. Table 3 presents several statistics of the OLS bias distribution for the 76 pharmaceutical companies using sample sizes from 5 to 15 years of monthly data.

The mean OLS bias of the entire sample of companies ranges from 0.17 in fifteen years to 0.27 in eight years (Table 3, column 2). The mean bias follows no clear pattern in samples ranging from five to nine years, but decreases from 0.21 to 0.17 in samples extended from ten to fifteen years. The median bias varies between 0.14 and 0.15 in most sample sizes. In general, the median provides a better representation of a central tendency in the OLS bias, especially in skewed samples like those representing the pharmaceutical industry.

We also observe that the standard deviation of the OLS bias gradually decreases as the sample size increases, starting at 0.492 in five-year samples and ending at 0.23 in fifteen-year samples. This finding confirms that OLS beta estimates from a larger sample are more stable. Further confirmation of the beneficial effect of increased sample size comes from the parallel decrease in the [Min Max] range of the OLS bias from [−0.72, 2.1] in five-year samples to [−0.29, 0.94] in fifteen-year samples.

Further analysis reveals that the 10th percentile of the OLS bias quickly decreases in absolute value from −0.27 in five-year samples to −0.16 in six-year samples and −0.12 in seven-year samples, stabilizing at a negligible bias between −0.07 and −0.08 in samples larger than seven years. The bias virtually disappears at the 25th percentile, assuming values between −0.03 and 0.03. In contrast, increases in the sample size appear to have a smaller relative impact on the 75th and 90th percentiles of the OLS bias distribution. In these percentiles the bias deviates closely around 0.3 and 0.6, respectively.

Figure 3 provides boxplots of the OLS bias for all sample sizes. The horizontal line inside each rectangle marks the median of the bias distribution. The lower and upper sides of the rectangle lying in the middle of each vertical plot mark the 25th and 75th percentiles of the distribution; the tips of the whiskers below and above the rectangle mark the 10th and 90th percentiles; and the dots below and above the whiskers stand for observations outside those percentiles.

Notice that the size and position of the rectangles and whiskers in Figure 3 mildly fluctuate in sample sizes between five and eight years. These features are remarkably stable in larger samples, a clear indication that the bias vanishes as the sample increases. For observations outside the 10th and 90th percentile, the OLS bias slightly decreases as the sample increases from five to nine years where it levels off. A significant positive OLS bias does not disappear in larger samples, possibly due to the fact that most outliers are large positive returns driven by good news. Strong signals of bad news may be rare since insiders have an incentive to release them late and in small doses, and more so the worse the news. Moreover, the rule of limited liability limits the scope of bad news by constraining the minimum share price at zero.

Table 4 lists the companies according to their OLS bias in a sample size of five years and reports the bias figures for five, eight, twelve, and fifteen years. In companies with a large positive bias in the initial five-year sample, the bias tends to decrease as the sample size increases, but remains positive and large. For example, the OLS bias of Poniard in the first five, eight, twelve, and fifteen years is 1.62, 1.24, 0.97, and 0.93, respectively. But in some companies the bias drops significantly. For example, the OLS bias of Genaera over the parallel periods is 2.1, 0.67, 0.35, and .34, respectively. With few exceptions, companies with a five-year OLS bias greater than 0.2 do not shed their large positive despite the extension of the sampling period (see bottom of Table 4).

Our results offer strong evidence that the impact of outlying observations on OLS-estimated beta tends to subside in a large sample. In a large majority of the companies studied, the OLS beta estimator is biased upwards. A positive bias in the five-year sample is likely to decrease by increased sample size, but remain economically significant. In contrast, only a small minority of the companies studied exhibit a negative OLS beta bias, which is likely to vanish in a large sample.

To further examine the relationship between the OLS bias and sample size, we run a simple regression of the bias of firm i in sample size S (years) against (S−5), the sample size in excess of five years. The following estimation model uses an interaction dummy variable to test for a differential impact of sample size on a negative and positive OLS bias:

$$\left(\beta_{OLS,i,S}-\beta_{k,i,S}\right)=\underset{(17.84)}{0.3734}-\underset{(-13.39)}{0.5384}\,D_{i,S}+\underset{(3.04)}{0.0215}\,D_{i,S}\left(S-5\right)-\underset{(-3.87)}{0.0137}\left(S-5\right)\quad R^2=0.318$$

Here the dummy variable $D_{i,s}$ = 1 stands for a negative bias and $D_{i,s}$ = 0 otherwise, S = 1, 2, …, 15 is the sample size in years, and i =1, 2, …, 76 identifies the company. Notice that all regression coefficients are statistically significant at the one percent level, indicating the presence of a systematic relationship between sample size and OLS beta bias. As indicated by the model's R-square value, 31.8% of the total variation in the sample is explained by this regression. The above regression further implies the following relationships in case of negative and positive bias:

$$\text{Negative bias: } \left(\beta_{OLS,i,S} - \beta_{k,i,S}\right) = -0.165 + 0.0078\left(S - 5\right)$$

$$\text{Positive bias: } \left(\beta_{OLS,i,S} - \beta_{k,i,S}\right) = 0.3734 - 0.0137\left(S - 5\right)$$

As the sample size increases, the OLS bias moves closer to zero – the negative bias increases and the positive bias decreases. In the sample containing five year of data, the average negative bias is −0.165 and the average positive bias is 0.3734; in the sample containing 15 years of data, the average negative bias is −0.087 and the average positive bias is 0.2364.

VI. Summary and Conclusions

We proposes a substitute for the OLS method in estimating the beta parameter, a key component of the estimated risk premium and overall cost of capital of the single company or project. Our proposed estimation method follows Huber's Robust M (HRM) methodology of mixed return, which makes our method resistant to return outliers – a known vulnerability of the OLS method. We undertake a battery of empirical tests comparing beta estimates of the two methods. To assess the adverse impact of outliers on the OLS-estimated beta of the single company, we use a sample of 76 Pharmaceutical companies to compare the two methods. Treating the difference between estimated OLS and HRM betas as an OLS bias, we show that the bias is negatively correlated with company size. Important company-specific events tend to have a larger impact on small companies as measured by the effects on the stock price, the

OLS-estimated beta, and the beta bias. Overall, the frequency of a positive OLS bias far exceeds that of a negative bias.

To determine whether the OLS bias negatively correlates with sample size and may be diminished by extending the sampling period beyond five years, we gradually increase that period to fifteen years. Overall, both negative and positive biases tend to diminish as the sample size increases. Despite this tendency, only negative biases are eliminated; positive biases remain economically significant.

Since the frequency of a positive OLS bias far exceeds that of a negative bias and cannot be eliminated by increasing the sample size, we conclude that a robust estimation method like HRM is preferable to the OLS to the extent that estimated parameters are expected to represent the majority of historical data.

We opened this paper with the statement that efficient estimation of the equity cost of public corporations is an essential component of calculating the required rate of return of real investment projects, and therefore the basis for a rational investment policy. Our evidence and that of other authors suggest that the original choice of the OLS as an estimation method for beta was a legitimate extrapolation of the effective role played by this method in the multi-asset context of portfolio selection. In that context, company-specific return volatility, including that of outliers, is largely diversified away even with limited diversification. Later contributions, like that of Blume (1975), added a cachet to the OLS and extended its life by improving its efficiency in the context of capital budgeting. In a world of uncertainty, a damage caused by an unknown beta bias will rarely expose the source of the damage. The relative performance of our model strongly suggests that robust estimators offer a promising alternative to the OLS in cost of capital estimation of the single company or project. The confirmed vulnerability of OLS beta estimates to the impact of outliers is a warning to those who must rely on a small sample.

Appendix: Statistical Methodology

Following Roll (1988), we model returns of individual stocks using a mixed return process with regular and outlier components.[6] This process allows for significant jumps (outliers) in the distribution of returns. The jump diffusion process, often used to price derivative assets, is the continuous analog of the mixed return process. Although the regular component distribution can be assumed normal, under the appropriate parameterization of the outlier component, the (overall) return can exhibit various levels of kurtosis and skewness. As such, the mixed return process is in line with empirical regularities in stock returns.

The general format of the mixed return process is:

$$R_{i,t} = K_{i,t} + d_{i,t} H_{i,t} \tag{1}$$

where $R_{i,t}$ is the return of stock i during period t, $K_{i,t}$ and $d_{i,t}$ $H_{i,t}$ are the regular and outlier components of company *i's* stock return, $d_{i,t}$ is a dichotomous (Bernouli) random variable that takes the value of unity with probability q when an outlier is present, and the value of zero with probability $(1–q)$ otherwise, $i = 1, 2,..., N$, $t =1, 2,..., T$, N is the number of stocks, and T is the number of historical return observations.
In the context of the general stock return model, the above mixed return model is written as:

$$R_{i,t} = \alpha_{k,i} + \beta_{k,i} R_{M,t} + e_{i,t} + d_{i,t} H_{i,t} \tag{2.a}$$

and

$$H_{i,t} = \alpha_{h,i} + \beta_{h,i} R_{M,t} + v_{i,t} \tag{2.b}$$

where $R_{M,,t}$ is the historical market return at time t, $\alpha_{k,i}$ is the intercept, $\beta_{k,i}$ is the beta coefficient of the company's regular return component, and $e_{i,t}$ is a white noise error term.

For the purpose of estimating the parameters of the regular return component, equation 2.a can be rewritten as

$$R_{i,t} = \alpha_{k,i} + \beta_{k,i} R_{M,t} + u_{i,t} \tag{3}$$

where the error term $u_{i,t} \equiv e_{i,t} + d_{i,t} H_{i,t}$ is decomposed into regular and outlier components using the rule:

[6] For the derivation of the model, see Theodossiou P. and A. Theodossiou, 2009, The Impact of Outliers on Stock Return Models: Implications for Event Studies and the Pricing of Risk. Working Paper.

$$e_{i,t}^{*}=\begin{cases} sign\left(u_{i,t}\right)c\sigma_{e,i} \text{ for} \left|u_{i,t}\right|>c\sigma_{e,i} \\ \\ u_{i,t} \text{ for} \left|u_{i,t}\right|\leq c\sigma_{e,i} \end{cases} \tag{4.a}$$

and

$$H_{i,t}^{*}=u_{i,t}-e_{i,t}^{*} \tag{4.b}$$

where $\sigma_{e,i}$ is the standard deviation of $e_{i,t}$, c is a predetermined constant, and *sign* takes the value of one for a positive $u_{i,t}$ and the value of minus one for a negative $u_{i,t}$. According to these equations, values of $u_{i,t}$ that fall outside the interval $[-c\sigma_{e,i}, c\sigma_{e,i}]$ are truncated to $-c\sigma_{e,i}$ or $c\sigma_{e,i}$ (depending on whether $u_{i,t}$ is negative or positive) and are assigned to $e^*_{i,t}$. The trimmed values of $u_{i,t}$ are assigned to $H_{i,t}^{*}$. Note that $e_{i,t}^{*}$ and $H_{i,t}^{*}$ are proxy variables for $e_{i,t}$ and $d_{i,t}H_{i,t}$, respectively.

Our estimation of the regular return regression is based on Huber's Robust M (HRM) estimator. Huber (1964, 1973) shows that when data follow a mixed probability distribution process, robust M estimators for the regression model can be obtained by minimizing the following maximum-likelihood type function:

$$\min_{\alpha_{k,i},\beta_{k,i}} S_i=\sum_{t=1}^{T}\rho\left(u_{i,t}\mid F_t\right) \tag{5.a}$$

where

$$\rho\left(u_{i,t}\mid F_t\right)=\begin{cases} 0.5u_{i,t}^2 \text{ for } \mid u_{i,t}\mid\leq c\sigma_{e,i} \\ \\ \left|u_{i,t}\right|c\sigma_{e,i}-0.5c^2\sigma_{e,i}^2 \text{ for } \left|u_{i,t}\right|>c\sigma_{e,i} \end{cases} \tag{5.b}$$

and

$$u_{i,t}=R_{i,t}-\left(\hat{\alpha}_{k,i}+\hat{\beta}_{k,i}F_t\right) \tag{5.c}$$

and other notations as defined above.[7] Note that the derivative of $\rho(u_{i,t}|F_t)$ with respect to $u_{i,t}$ is

$$\partial\rho\left(u_{i,t}\middle|F_t\right)/\partial u_{i,t}=e_{i,t}^{*} \tag{6}$$

[7]A review of the estimation and statistical properties of Huber's method can be found in Judge et al. (1985), Ch. 20. pp. 828-834; see also Huber (1981). For a thorough analysis of robust estimation methods, see Hampel et al. (1986). An extensive review on outliers in statistics can be found in Beckman and Cook (1983).

where $e_{i,t}^{*}$ is as defined above. The HRM estimators for the regression intercept and slope satisfy the following normal equations (first order conditions for minimization),

$$\frac{\partial S_i}{\partial \alpha_{k,i}} = \sum_{t=1}^{T} \hat{e}_{i,t} = 0 \tag{7.a}$$

and

$$\frac{\partial S_i}{\partial \beta_{k,i}} = \sum_{t=1}^{T} F_t' \, \hat{e}_{i,t} = \underset{\sim}{o} \tag{7.b}$$

These equations are solved iteratively using the algorithm

$$\hat{\theta}_i^{s+1} = \hat{\theta}_i^{s} + \left(\sum_{t=1}^{T} Z_t' Z_t \right)^{-1} \sum_{t=1}^{T} Z_t' \hat{e}_{i,t} \tag{8.a}$$

where

$$\hat{e}_{i,t} = \begin{cases} \hat{u}_{i,t} = Y_t - Z_t \hat{\theta}_i \text{ for } |\, \hat{u}_{i,t} \,| < c\hat{\sigma}_{e,i} \\ \\ sign\left(\hat{u}_{i,t}\right) c\hat{\sigma}_{e,i} \text{ for } |\, \hat{u}_{i,t} \,| \geq c\hat{\sigma}_{e,i} \end{cases} \tag{8.b}$$

and where $\theta_i' = \left[\alpha_{k,i}, \beta_{k,i}'\right], Z_t = \left[1, F_t\right]$, $\hat{\sigma}_{e,i}$ is the HRM estimator for $\sigma_{e,i}^2$, and s is the iteration counter. OLS estimates of θ can be used as starting values for the algorithm. The algorithm converges in 5 to 20 iterations to an accuracy of 10^{-5}. The regression variance is estimated by

$$\hat{\sigma}_{e,i}^2 = \left(\frac{g_i^2}{T-2} \right) \sum_{t=1}^{T} \hat{e}_{i,t}^2 \tag{9.a}$$

where

$$g_i = \left[1 + \frac{2}{T} \left(\frac{1-m_i}{m_i} \right) \right] \frac{1}{m_i} \tag{9.b}$$

and m_i is the proportion of untrimmed returns of company i. Standard errors for the estimator are computed from

$$\operatorname{var}\left(\hat{\theta}_i\right) = \hat{\sigma}_{e,i}^2 \left(\sum_{t=1}^{T} Z_t' Z_t \right)^{-1} \tag{10}$$

Excess or abnormal returns of individual stocks are computed from the data using the equation

$$\hat{u}_{i,t} = R_{i,t} - \left(\hat{\alpha}_{k,i} + \hat{\beta}_{k,i} F_t\right) \tag{11}$$

and outlier excess returns from

$$\hat{H}_{i,t} = \hat{u}_{i,t} - \hat{e}_{i,t} \tag{12}$$

Most $\hat{H}_{i,t}$ values are zero. Large positive or negative values are indicative of outlier returns. Unlike $\hat{e}_{i,t}$, the residual $\hat{u}_{i,t}$ captures the full impact of outlier returns.

References

Anderson, T. W. 1984. *An Introduction to Multivariate Analysis*, 2nd Ed. John Wiley & Sons, Inc.

Beckman, R. J. and R. D. Cook. 1983. Outliers. *Technometrics* 25(2):119-149.

Blume, M. E. 1975. Betas and Their Regression Tendencies. *Journal of Finance* 30:785-799.

Bruner, R.F., K. M. Eades, R. Harris, and R. C. Higgins. 1998. Best Practices in Estimating the Cost of Capital: Survey and Synthesis. *Financial Practice and Education* 8:13-28.

Chan, L.K.C. and J. Lakonishok. 1992. Robust Measurement of Beta Risk. *Journal of Financial and Quantitative Analysis* 27: 265-282.

Fama, E. F. and K. R. French. 1992. The Cross-Section of Expected Stock Returns. *Journal of Finance* 47:427-466.

Fama, E. F., and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 25:23-49.

Genton, M. G., and E. Ronchetti. 2008. Robust Prediction of Beta. *Computational Methods in Financial Engineering, Essays in Honour of Manfred Gilli*, Springer, 147-161.

Graham, J. R. and C. R. Harvey. 2001. The Theory and Practice of Corporate Finance: Evidence from the Field. *Journal of Financial Economics* 60: 187-243.

Gray, S., J. Hall, J. Bowman, T. Brailsford, R. Faff, and B. Officer. May 2005. *The Performance of Alternative Techniques for Estimating Equity Betas of Australian Firms*. Report Prepared for the Energy Networks Association, 1-41.

Hampel, F.R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. 1986. *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

Huber, P.J. 1981. *Robust Statistics*. John Wiley & Sons, Inc.

Huber, P.J. Mar. 1964. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics* 35(5): 73-101.

Huber, P.J. Sep. 1973. Robust Regression Asymptotics, Conjectures and Monte Carlo. *The Annals of Statistics* 1(5): 799-821.

Judge, G.J., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T-C. Lee. 1985. *The Theory and Practice of Econometrics*, 2nd Ed. John Wiley & Sons, Inc.

Keehan S., A. Sisko, C. Truffer, S. Smith, C. Cowan, J. Poisal, M. Kent Clemens the National Health Expenditure Accounts Projections Team. 2008. Health Spending Projections Through 2017: The Baby-Boom Generation Is Coming To Medicare. *Health Affairs* 27(2):w145-w155.

Knez, P. J., and M. J. Ready. 1997. On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions. *Journal of Finance* 52(4):1355-1382.

Martin, R. D. and T. Simin. 2003. Outlier Resistant Estimates of Beta. *Financial Analysts Journal* 59, 56–69.

McDonald, J. B., R.A. Michelfelder, and P. Theodossiou. 2008. Evaluation of Robust Regression Estimation Methods and Intercept Bias: A Capital Asset Pricing Model Application, Working Paper, Rutgers University.

Roll, R. 1988. R2. *Journal of Finance*, 43(3): 541-566.

Theodossiou P. and A. Theodossiou. 2009. Impact of Outliers on Stock Return Models. Implications for Event Studies and the Pricing of Risk. *Working Paper*.

Figure 1. Lorenz Curve for Market Capitalization

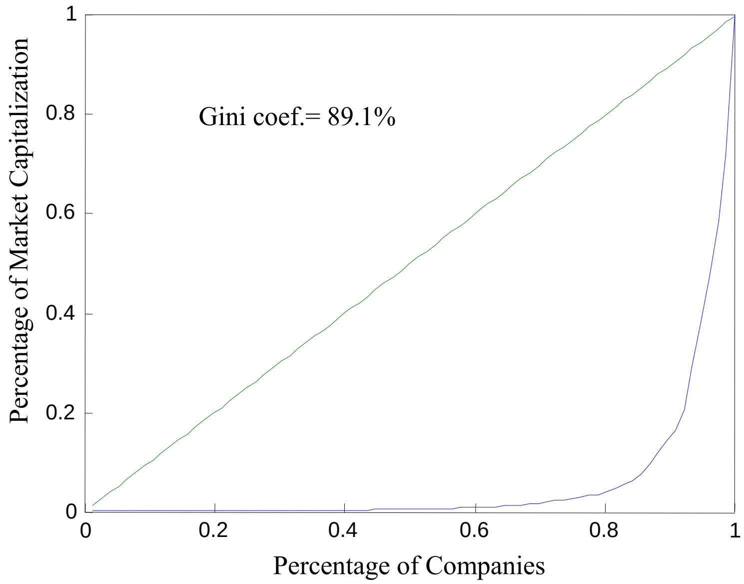

Figure 2A. Genaera Corp

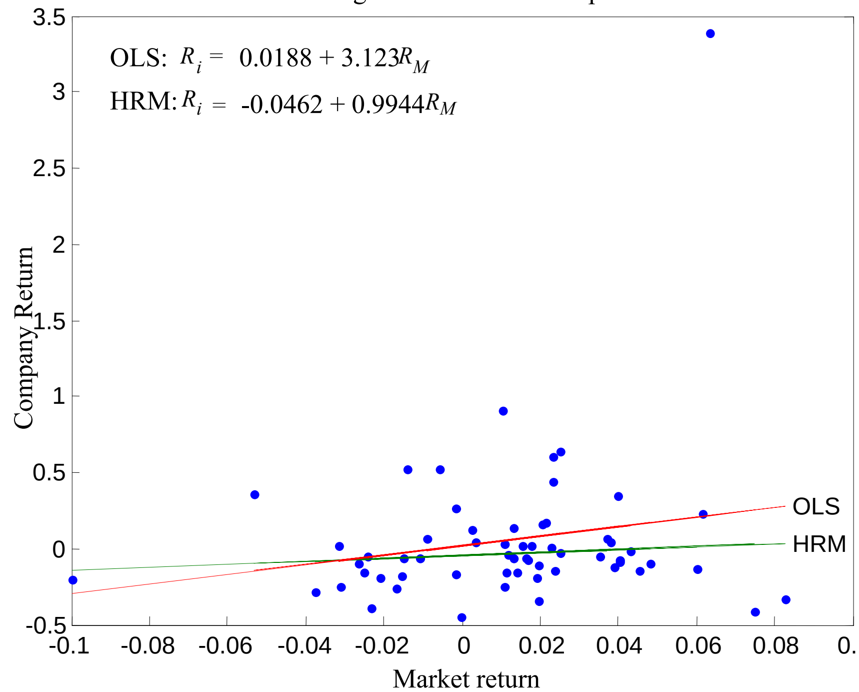

Figure 2B. Amylin Pharmaceuticals Inc

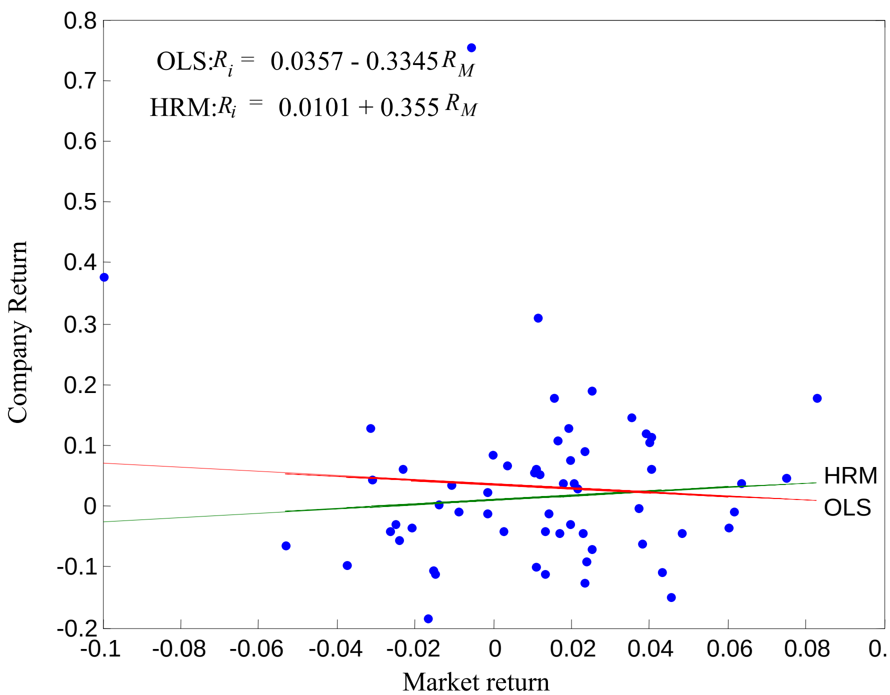

Figure 2C. Poniard Pharmaceuticals Inc

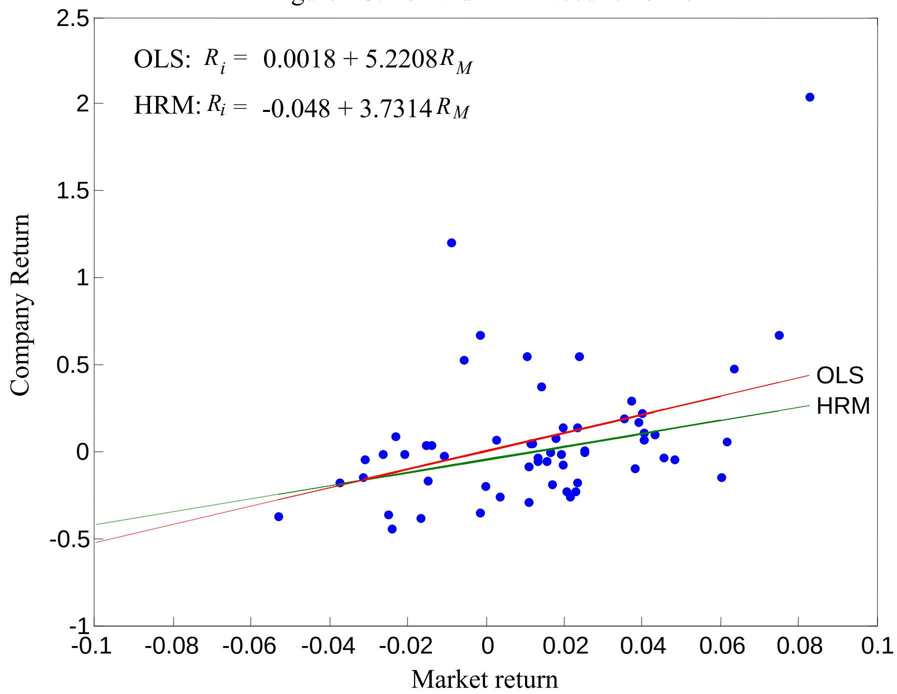

Figure 2D. Merck & Co

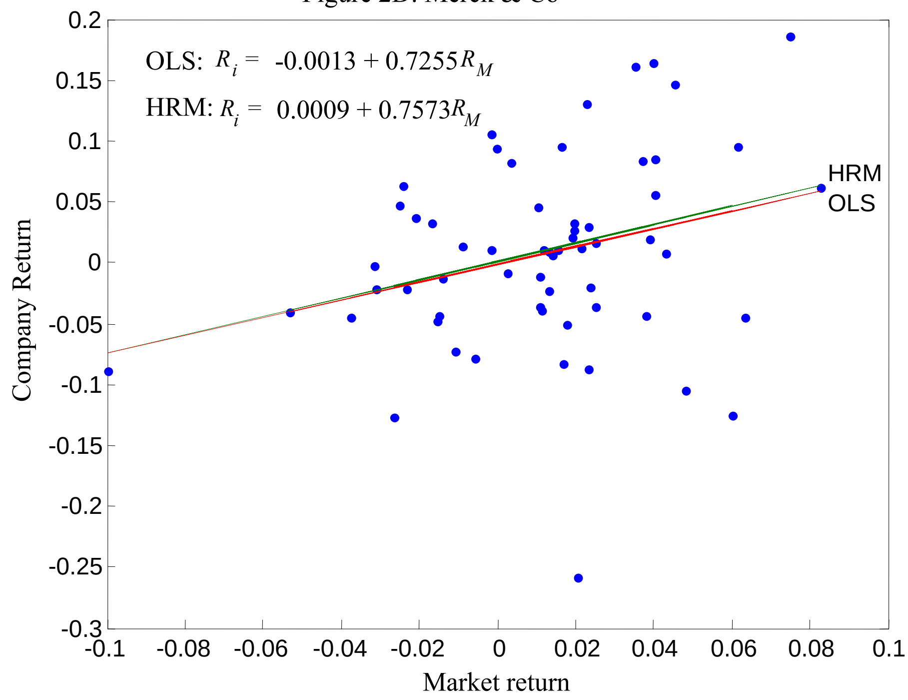

Figure 3. Boxplots for OLS Bias

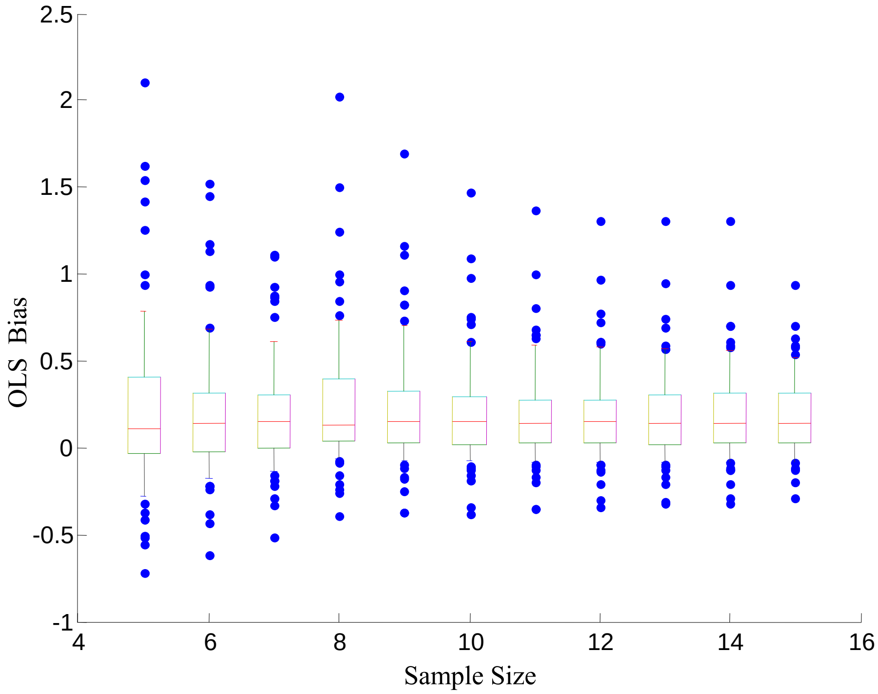

TABLE 1. Pharmaceutical Company Names, Market Capitalization, and Betas: September 2002 - August 2007

No	TICKER	Company Name	Median Capital. (MM-$)	Cusum1 Weight (%)	Cusum2 Weight (%)	Mean Return (%)	SD of Return (%)	β_{OLS}	β_k	$\lvert\beta_{OLS}-\beta_k\rvert$
1	HCTL	Healthcare Technologies Ltd	6	100.0	0.001	6.4	37.3	2.7769	1.7812	0.9957
2	ICCC	Immucell Corp	13	100.0	0.003	1.9	14.9	2.1047	1.5797	0.5250
3	CYAN	Cyanotech Corp	14	100.0	0.005	0.5	21.3	0.4492	0.3251	0.1241
4	IG	IGI Inc	15	100.0	0.007	2.8	22.4	1.3262	0.6024	0.7238
5	POLXF	Polydex Pharmaceuticals Ltd	18	100.0	0.009	1.6	22.1	1.5412	1.0981	0.4432
6	NXXI	Nutrition 21 Inc	36	100.0	0.014	5.6	34.5	1.9308	1.5802	0.3506
7	RPRX	Repros Therapeutics Inc	37	100.0	0.020	5.7	23.5	1.5258	1.7938	0.2680
8	QSC	Questcor Pharmaceuticals Inc	39	100.0	0.025	0.4	18.6	0.7536	0.6610	0.0926
9	UG	United-Guardian Inc	39	100.0	0.031	2.3	11.1	1.2826	1.0758	0.2067
10	ALSE	Alseres Pharmaceuticals Inc	40	100.0	0.036	0.4	19.3	1.6991	1.4124	0.2867
11	ISV	Insite Vision Inc	41	100.0	0.042	2.2	22.7	0.8840	0.3231	0.5609
12	MZT	Matritech Inc	48	100.0	0.049	-2.4	20.0	0.8241	1.0485	0.2244
13	NAII	Natural Alternatives	50	100.0	0.056	2.1	11.2	0.0652	0.1413	0.0761
14	PARS	Pharmos Corp	56	99.9	0.064	-0.9	21.4	2.1923	1.6831	0.5092
15	CYTR	Cytrx Corp	64	99.9	0.073	6.2	27.8	2.9890	1.5720	1.4169
16	GNLB	Genelabs Technologies Inc	67	99.9	0.083	0.5	26.1	2.6859	2.7768	0.0909
17	PARD	Poniard Pharmaceuticals Inc	74	99.9	0.093	6.2	39.6	5.2208	3.5960	1.6247
18	CYTO	Cytogen Corp	75	99.9	0.104	-1.0	21.7	4.2286	4.1560	0.0726
19	GTCB	GTC Biotherapeutics Inc	76	99.9	0.114	2.3	26.4	2.1335	0.8793	1.2542
20	RGEN	Repligen Corp	94	99.9	0.128	2.4	18.5	1.7898	1.3806	0.4092
21	GENR	Genaera Corp	97	99.9	0.141	5.5	51.3	3.1230	1.0204	2.1027
22	EMIS	Emisphere Technologies Inc	98	99.9	0.155	2.6	21.5	0.8067	1.3125	0.5058
23	BIOM	Biomira Inc	104	99.8	0.170	1.2	23.1	3.7484	2.2006	1.5478
24	DUSA	Dusa Pharmaceuticals Inc	108	99.8	0.185	1.7	17.5	1.7086	1.6635	0.0451
25	OXGN	Oxigene Inc	109	99.8	0.201	4.5	28.8	2.5323	1.8793	0.6531
26	AMAG	Advanced Magnetics Inc	110	99.8	0.216	7.2	25.6	1.0763	0.9708	0.1055
27	AVAN	Avant Immunotherapeutics Inc	120	99.8	0.233	0.2	17.9	2.4267	2.0264	0.4003
28	OSCI	Oscient Pharmaceuticals Corp	121	99.8	0.250	-0.7	21.0	2.9262	2.5453	0.3809
29	VICL	Vical Inc	121	99.7	0.268	1.4	17.6	3.2535	2.7228	0.5307
30	ANIK	Anika Therapeutics Inc	123	99.7	0.285	5.9	17.4	0.9971	0.7762	0.2209
31	ZILA	Zila Incorporated	149	99.7	0.306	0.5	16.0	1.8360	1.3733	0.4627

32	SCLN	Sciclone Pharmaceuticals Inc	154	99.7	0.328	1.5	21.3	1.5470	1.6080	0.0610
33	NEOG	Neogen Corp	154	99.7	0.350	2.2	8.4	0.7927	0.6926	0.1001
34	STEM	Stemcells Inc	174	99.7	0.374	4.2	30.7	2.4988	1.5640	0.9348
35	IMMU	Immunomedics Inc	183	99.6	0.400	0.9	22.9	3.0021	3.0370	0.0349
36	GNTA	Genta Inc	202	99.6	0.429	-2.7	26.3	2.6472	2.3330	0.3141
37	IMGN	Immunogen Inc	207	99.6	0.458	1.8	16.3	2.3985	2.8091	0.4107
38	NRGN	Neurogen Corp	221	99.5	0.489	-0.1	13.2	1.6421	1.4027	0.2394
39	BNT	Bentley Pharmaceuticals	239	99.5	0.523	1.6	16.3	2.0541	1.8069	0.2472
40	QDEL	Quidel Corp	242	99.5	0.558	3.6	18.2	2.2244	2.2482	0.0237
41	VIVO	Meridian Bioscience Inc	261	99.4	0.595	4.4	10.3	1.2579	1.1173	0.1407
42	XOMA	Xoma Ltd	265	99.4	0.632	0.4	17.4	1.7119	2.2271	0.5152
43	IDEV	Indevus Pharmaceuticals Inc	279	99.4	0.672	3.9	17.8	1.7722	1.4166	0.3556
44	CEGE	Cell Genesys Inc	313	99.3	0.716	-0.9	14.0	1.7983	2.0072	0.2089
45	ISIS	Isis Pharmaceuticals Inc	367	99.3	0.768	1.7	17.3	2.2863	2.1267	0.1596
46	NOVN	Noven Pharmaceuticals Inc	413	99.2	0.827	1.2	13.9	0.0813	0.4524	0.3711
47	ENZN	Enzon Pharmaceuticals Inc	458	99.2	0.891	-1.0	11.2	1.1700	1.0761	0.0939
48	CHTT	Chattem Inc	650	99.1	0.984	2.5	10.4	0.3589	0.3091	0.0498
49	REGN	Regeneron Pharmaceut	739	99.0	1.088	2.8	23.7	2.3746	1.8622	0.5124
50	MEDX	Medarex Inc	922	98.9	1.219	3.2	17.7	3.7557	3.5675	0.1882
51	PRX	Par Pharmaceutical Cos Inc	1,068	98.8	1.370	0.4	11.6	0.2900	0.2910	0.0010
52	BLUD	Immucor Inc	1,097	98.6	1.526	4.5	12.2	0.9333	1.0415	0.1081
53	VRTX	Vertex Pharmaceuticals Inc	1,309	98.5	1.711	2.1	14.1	1.6295	1.5072	0.1223
54	PRGO	Perrigo Co	1,363	98.3	1.904	1.5	8.1	0.7589	0.5705	0.1884
55	VRX	Valeant Pharmaceuticals Intl	1,649	98.1	2.137	1.7	14.2	1.6219	0.9932	0.6286
56	TECH	Techne Corp	1,670	97.9	2.374	1.6	7.8	0.7158	1.0354	0.3196
57	NTY	Nbty Inc	1,689	97.6	2.613	2.4	14.1	1.0672	1.1178	0.0506
58	OSIP	Osi Pharmaceuticals Inc	1,869	97.4	2.878	2.8	19.1	0.8010	1.3575	0.5565
59	IDXX	Idexx Labs Inc	2,003	97.1	3.162	2.5	6.9	0.0347	-0.1497	0.1844
60	AMLN	Amylin Pharmaceuticals Inc	2,285	96.8	3.485	3.2	14.2	-0.3345	0.3864	0.7208
61	WPI	Watson Pharmaceuticals Inc	3,246	96.5	3.945	0.8	8.6	0.9784	1.0080	0.0296
62	MYL	Mylan Laboratories Inc	4,687	96.1	4.609	0.5	8.0	0.1646	0.0721	0.0925
63	SEPR	Sepracor Inc	4,846	95.4	5.295	4.1	17.8	1.7365	1.2022	0.5343
64	BRL	Barr Pharmaceuticals Inc	5,047	94.7	6.010	1.2	8.9	0.3774	0.3986	0.0212
65	AGN	Allergan Inc	11,503	94.0	7.639	1.4	5.7	0.7946	0.7518	0.0428
66	GENZ	Genzyme Corp	14,484	92.4	9.691	2.2	8.7	1.1892	0.9108	0.2785
67	BIIB	Biogen Idec Inc	15,186	90.3	11.84	1.4	10.4	0.2627	0.4247	0.1620

68	GILD	Gilead Sciences Inc	16,143	88.2	14.13	2.8	7.4	0.6900	0.6373	0.0526
69	FRX	Forest Laboratories	16,290	85.9	16.44	0.4	8.9	0.6633	0.9229	0.2596
70	SGP	Schering-Plough	28,276	83.6	20.44	0.8	7.1	0.6600	0.5821	0.0779
71	WYE	Wyeth	59,162	79.6	28.82	0.6	6.8	1.0676	0.8335	0.2341
72	LLY	Lilly (Eli) & Co	64,191	71.2	37.91	0.4	6.8	0.8103	0.7584	0.0519
73	ABT	Abbott Laboratories	67,663	62.1	47.49	0.9	5.1	0.3556	0.3884	0.0329
74	AMGN	Amgen Inc	77,932	52.5	58.53	0.4	7.1	0.7282	0.6569	0.0714
75	MRK	Merck & Co	96,892	41.5	72.26	0.7	7.9	0.7255	0.7675	0.0421
76	PFE	Pfizer Inc	195,878	27.7	100.0	-0.1	5.7	0.7324	0.7258	0.0066

Descriptive Statistics	Median Capital.	Cusum1 Weight	Cusum2 Weight	Mean Return	SD of Return	β_{OLS}	β_k	$\lvert\beta_{OLS}-\beta_k\rvert$
Mean(X)	9,290	1.316	6.367	1.9	17.1	1.5482	1.3265	0.3558
SD(X)	28,631	4.055	16.907	2.0	8.6	1.0691	0.8641	0.4045
Minimum	6	0.0008	0.0008	-2.6713	5.1015	-0.3345	-0.1497	0.0010
10 Percentile	39	0.01	0.03	-0.13	7.13	0.3556	0.3864	0.0421
25 Percentile	76	0.01	0.11	0.47	10.27	0.7536	0.6926	0.0779
50 Percentile	230	0.03	0.51	1.66	17.33	1.426	1.1077	0.2292
75 Percentile	1,869	0.26	2.88	2.81	21.68	2.2244	1.7938	0.5092
90 Percentile	16,290	2.31	16.44	5.50	26.37	2.989	2.5453	0.7238
Maximum	195,878	27.7	100	7.17	51.33	5.2208	4.156	2.1027

Notes: The statistics above are based on data over the sampling period 9/2002-8/2007. Companies are sorted in ascending order of market capitalization. Median Capital is the median value of market capitalization over the sampling period. The overall market capitalization of the 76 companies is $706.032 billion. Cusum1 weight is the percent of cumulative capitalization out of total capitalization starting from the largest (bottom of the table) to the smallest. Cusum2 weight is the percent of cumulative capitalization to total capitalization from the smallest (top of the table) to the largest. Mean and SD refer to the deviation of monthly returns computed over the sampling period. Coefficients β_{OLS} and β_k are the OLS and HRM estimators for each firm over the sampling period and $\lvert\beta_{OLS}-\beta_k\rvert$ is their absolute difference.

TABLE 2. Market Capitalization, Stock Price, Volume, and Returns during Extreme Periods

Date	Market Capitalization (MM-$)	Stock Price ($)	Shares Outstanding (MM)	Volume (MM)	Company Return (%)	Market Return (%)
Genaera Corp						
2002/10	13.6	0.38	35.7	4.5	-41.5	7.5
2002/12	22.8	0.64	35.7	5.5	36.2	-5.3
2003/01	13.9	0.39	35.7	2.2	-39.1	-2.3
2003/03	21.8	0.61	35.7	4.4	90.6	1.0
2003/04	14.6	0.41	35.7	26.8	-32.8	8.3
2003/05	64.2	1.80	35.7	46.0	339.0	6.4
2003/07	98.5	2.72	36.2	18.2	61.0	2.3
2003/08	161.2	4.45	36.2	30.1	63.6	2.5
2004/01	221.0	4.70	47.0	28.0	43.7	2.3
2004/07	157.8	3.00	52.6	8.0	-28.6	-3.8
2005/08	145.3	2.54	57.2	23.9	52.1	-0.6
2006/06	57.5	0.55	104.6	14.1	-44.7	0.0
2007/01	26.2	0.25	104.7	38.2	-34.2	1.9
2007/02	39.8	0.38	104.7	30.6	52.1	-1.4
2007/04	60.7	0.58	104.7	25.1	35.0	4.0
Amylin Pharmaceuticals Inc						
2002/09	1,359.2	16.62	81.8	22.0	37.8	-10.0
2005/06	2,183.7	20.93	104.3	111.7	31.0	1.2
2005/08	3,595.3	32.75	109.8	123.0	75.5	-0.6
Poniard Pharmaceuticals Inc (Old NeoRx)						
200210	17.42	0.65	26.8	2.2	66.9	7.5
200303	20.40	0.76	26.8	0.9	55.1	1.0
200304	62.01	2.31	26.8	11.7	204.0	8.3
200309	171.4	6.19	27.7	10.3	120.3	-0.9
200508	34.57	1.01	34.2	6.1	53.0	-0.6
200602	46.09	1.34	34.3	5.8	67.8	-0.2
200611	159.9	7.01	22.8	2.0	55.1	2.4
Merck & Co Inc						
2003/10	99,068.5	44.25	2238.8	203.8	-12.6	6.0
2004/09	73,217.3	33.00	2218.7	254.8	-25.8	2.1
2004/11	62,136.7	28.02	2217.6	578.4	-10.5	4.8

Notes: MM stands for millions. Market capitalization is equal to price times the number of shares outstanding.

TABLE 3. Statistics on the Differences of OLS and HRM Betas Estimators

Years	OBS	Mean	SDEV	Min	10%	25%	50%	75%	90%	Max
5	76	0.22	0.49	-0.72	-0.27	-0.03	0.11	0.41	0.72	2.10
6	76	0.20	0.38	-0.62	-0.16	-0.03	0.14	0.31	0.68	1.52
7	76	0.19	0.31	-0.52	-0.12	-0.01	0.15	0.31	0.55	1.12
8	76	0.27	0.40	-0.39	-0.07	0.03	0.13	0.39	0.73	2.03
9	76	0.23	0.35	-0.37	-0.07	0.03	0.15	0.32	0.69	1.70
10	76	0.21	0.31	-0.38	-0.07	0.01	0.15	0.31	0.60	1.47
11	76	0.19	0.28	-0.35	-0.07	0.02	0.14	0.28	0.58	1.37
12	76	0.18	0.27	-0.34	-0.08	0.02	0.15	0.28	0.57	1.31
13	74	0.19	0.27	-0.32	-0.07	0.02	0.14	0.31	0.57	1.31
14	74	0.19	0.27	-0.32	-0.08	0.03	0.14	0.31	0.54	1.31
15	66	0.17	0.23	-0.29	-0.08	0.03	0.14	0.31	0.51	0.94

TABLE 4. Sample Size and OLS Beta Bias: Comparing OLS and HRM

NO	TICKER	Company Name	5 Years	8 Years	12 Years	15 Years
1	GENR	Genaera Corp	2.1027	0.6707	0.3542	0.344
2	PARD	Poniard Pharmaceuticals Inc	1.6247	1.2450	0.9705	0.9356
3	BIOM	Biomira Inc	1.5478	0.5660	0.3997	0.4021
4	CYTR	Cytrx Corp	1.4169	0.6990	0.4027	0.359
5	GTCB	GTC Biotherapeutics Inc	1.2542	0.8446	0.7232	n/a
6	HCTL	Healthcare Technologies Ltd	0.9957	0.6873	0.4623	0.4161
7	STEM	Stemcells Inc	0.9348	0.6206	0.2520	0.1969
8	IG	IGI Inc	0.7238	0.4252	0.2217	0.2459
9	OXGN	Oxigene Inc	0.6531	0.2696	0.1508	n/a
10	VRX	Valeant Pharmaceuticals Intl	0.6286	0.3724	0.3551	0.2919
11	ISV	Insite Vision Inc	0.5609	0.2434	0.1667	n/a
12	SEPR	Sepracor Inc	0.5343	0.0823	0.1458	0.1217
13	VICL	Vical Inc	0.5307	0.1117	0.1476	n/a
14	ICCC	Immucell Corp	0.5250	0.3338	0.2424	0.2683
15	REGN	Regeneron Pharmaceut	0.5124	0.7625	0.5256	0.6296
16	PARS	Pharmos Corp	0.5092	0.3015	0.0326	0.1444
17	ZILA	Zila Incorporated	0.4627	0.4074	0.3076	0.2883
18	POLXF	Polydex Pharmaceuticals Ltd	0.4432	0.3478	0.3522	0.3603
19	RGEN	Repligen Corp	0.4092	-0.0435	0.1205	0.1558
20	AVAN	Avant Immunotherapeutics Inc	0.4003	0.3231	0.2321	n/a
21	OSCI	Oscient Pharmaceuticals Corp	0.3809	1.5001	0.7712	0.6998
22	IDEV	Indevus Pharmaceuticals Inc	0.3556	0.9623	0.6043	0.5404
23	NXXI	Nutrition 21 Inc	0.3506	0.3946	0.1248	0.1968
24	GNTA	Genta Inc	0.3141	0.2292	0.2830	0.3307
25	ALSE	Alseres Pharmaceuticals Inc	0.2867	0.2088	0.1774	0.314
26	GENZ	Genzyme Corp	0.2785	0.1172	0.0243	n/a
27	BNT	Bentley Pharmaceuticals	0.2472	0.1073	0.1159	n/a
28	NRGN	Neurogen Corp	0.2394	0.1399	0.1716	0.1883
29	WYE	Wyeth	0.2341	0.1107	0.0665	0.0605
30	ANIK	Anika Therapeutics Inc	0.2209	0.4184	0.1842	n/a
31	UG	United-Guardian Inc	0.2067	0.1112	0.1962	0.1539
32	PRGO	Perrigo Co	0.1884	0.1083	0.0183	0.0147
33	MEDX	Medarex Inc	0.1882	0.9956	0.5737	0.5844
34	IDXX	Idexx Labs Inc	0.1844	0.1551	0.2572	0.18
35	ISIS	Isis Pharmaceuticals Inc	0.1596	-0.0851	-0.0573	0.0278
36	VIVO	Meridian Bioscience Inc	0.1407	0.2822	0.0916	0.1002
37	CYAN	Cyanotech Corp	0.1241	0.7274	0.3764	0.3992
38	VRTX	Vertex Pharmaceuticals Inc	0.1223	0.1921	0.1085	0.1021
39	AMAG	Advanced Magnetics Inc	0.1055	-0.0575	-0.0955	-0.0882
40	NEOG	Neogen Corp	0.1001	0.1187	0.0844	0.1023
41	ENZN	Enzon Pharmaceuticals Inc	0.0939	0.0330	0.1632	0.1962
42	QSC	Questcor Pharmaceuticals Inc	0.0926	-0.2545	-0.0920	-0.1105
43	MYL	Mylan Laboratories Inc	0.0925	0.1059	0.0218	0.0303
44	SGP	Schering-Plough	0.0779	-0.0377	-0.0786	-0.0782
45	CYTO	Cytogen Corp	0.0726	0.6179	0.6154	0.5905
46	AMGN	Amgen Inc	0.0714	0.0265	0.0401	0.0311
47	GILD	Gilead Sciences Inc	0.0526	0.2049	0.1522	0.0975

48	LLY	Lilly (Eli) & Co	0.0519	-0.0561	-0.0542	-0.0553
49	CHTT	Chattem Inc	0.0498	0.1190	0.2317	0.2064
50	DUSA	Dusa Pharmaceuticals Inc	0.0451	0.1350	0.1984	0.2518
51	AGN	Allergan Inc	0.0428	-0.1597	-0.0555	-0.0696
52	PFE	Pfizer Inc	0.0066	-0.0701	-0.1266	-0.1274
53	PRX	Par Pharmaceutical Cos Inc	-0.0010	0.2029	0.2641	0.3255
54	BRL	Barr Pharmaceuticals Inc	-0.0212	0.0647	0.0788	0.0706
55	QDEL	Quidel Corp	-0.0237	0.0831	0.0357	0.0240
56	WPI	Watson Pharmaceuticals Inc	-0.0296	-0.0309	0.0350	0.0718
57	ABT	Abbott Laboratories	-0.0329	0.0475	-0.0346	-0.0337
58	IMMU	Immunomedics Inc	-0.0349	2.0272	1.3122	n/a
59	MRK	Merck & Co	-0.0421	-0.0447	-0.0670	-0.0595
60	NTY	Nbty Inc	-0.0506	-0.2091	-0.1369	0.0297
61	SCLN	Sciclone Pharmaceuticals Inc	-0.0610	0.6698	0.4606	0.4238
62	NAII	Natural Alternatives	-0.0761	0.0909	0.1666	0.1039
63	GNLB	Genelabs Technologies Inc	-0.0909	-0.2359	-0.2076	-0.1955
64	BLUD	Immucor Inc	-0.1081	-0.3906	-0.3041	-0.2877
65	BIIB	Biogen Idec Inc	-0.1620	0.0515	0.0509	0.0573
66	CEGE	Cell Genesys Inc	-0.2089	0.2364	0.0418	0.0288
67	MZT	Matritech Inc	-0.2244	0.3905	0.6092	0.506
68	FRX	Forest Laboratories	-0.2596	0.0142	-0.0311	-0.021
69	RPRX	Repros Therapeutics Inc	-0.2680	0.0284	0.0741	n/a
70	TECH	Techne Corp	-0.3196	0.0913	-0.0201	-0.0334
71	NOVN	Noven Pharmaceuticals Inc	-0.3711	0.0836	0.1094	0.1393
72	IMGN	Immunogen Inc	-0.4107	-0.0661	0.0390	0.0417
73	EMIS	Emisphere Technologies Inc	-0.5058	0.0049	0.0400	0.148
74	XOMA	Xoma Ltd	-0.5152	0.0661	-0.0327	-0.0755
75	OSIP	Osi Pharmaceuticals Inc	-0.5565	0.2994	0.2126	0.1437
76	AMLN	Amylin Pharmaceuticals Inc	-0.7208	0.0856	-0.3366	-0.285

Descriptive Statistics of the OLS Bias	5 Years	8 Years	12 Years	15 Years
Mean(X)	0.2218	0.2662	0.1844	0.1690
SD(X)	0.4921	0.3965	0.2732	0.2339
Minimum	-0.7208	-0.3906	-0.3366	-0.2877
10% Percentile	-0.2680	-0.0661	-0.0786	-0.0782
25% Percentile	-0.0349	0.0330	0.0243	0.0278
50% Percentile	0.1139	0.1270	0.1467	0.1415
75% Percentile	0.4092	0.3946	0.2830	0.3140
90% Percentile	0.7238	0.7274	0.5737	0.5060
Maximum	2.1027	2.0272	1.3122	0.9356

Notes: Companies are sorted based on the OLS and HRM difference in the 5-year sample. This table is intended to show the impact of larger samples on the OLS bias.