Assessing the reliability of regression-based estimates of risk

17 June 2013

Neil Diamond and Robert Brooks, Monash University
Stephen Gray and Jason Hall, SFG Consulting

Contents

1. Preparation of this report

This report was prepared by Professor Stephen Gray, Dr Jason Hall, Professor Robert Brooks and Dr Neil Diamond. Professor Gray, Dr Hall, Professor Brooks and Dr Diamond acknowledge that they have read, understood and complied with the Federal Court of Australia's Practice Note CM 7, Expert Witnesses in Proceedings in the Federal Court of Australia. Professor Gray, Dr Hall, Professor Brooks and Dr Diamond provide advice on cost of capital issues for a number of entities but have no current or future potential conflicts.

2. Executive summary

In prior decisions, the Australian Energy Regulator ("the AER" or "the regulator") has estimated the systematic risk of the benchmark firm by relying almost exclusively on regression analysis of stock returns, performed on a small sample of nine Australian-listed stocks. Only five of these remain listed and at no time have all nine stocks had financial information available for analysis. The regulator has not placed material reliance on a larger sample of stocks listed in the United States.

This presents the regulator with a small sample problem. The problem is that we have no way of knowing what the risk estimate of comparator firms would have been if a greater number of firms had been available for analysis. There is the potential that the risk estimates from this limited sample do not represent a reasonable estimate of risk for the benchmark firm, and could have been markedly different if we had observed returns for a different set of nine comparable firms, or a different set of outcomes for the same nine firms.

In this paper we measure the relationship between sample size and reliability, in the following two ways, considering both beta estimates from the Sharpe-Lintner and Black Capital Asset Pricing Models ("CAPMs"), and risk premium estimates from the Fama & French 3-factor model.

1. We measure the dispersion of risk estimates across different samples of firms from the same industry. This measures how different the estimate of risk would be purely from selection of a different set of firms of the same sample size. The greater this dispersion, the less reliable the risk estimates. We document that the dispersion of risk estimates is reduced by about 30% if sample size is increased from nine to 18 firms, and by about 50% if sample size is increased to 27 firms.

2. We measure the variation in risk estimates over time for the same sample of firms. For estimating the cost of capital it is important that, if the risk estimate is based upon historical data, it is reasonably stable over time. This holds even if the true risk is unchanged from one period to another and the regression-based estimate of risk precisely measures this risk. The reason for this is that the regression estimate is based entirely on historical data, so if there is a large degree of variation in the estimate over time (even if this does indeed measure the true risk), then the risk estimate will be of little use in predicting future risk. We document substantial variation in risk estimates over time, even for samples much larger than nine firms and even for sub-samples in which the initial risk estimates were closest amongst sample firms.

There are three clear implications of our findings:

- Exclusive reliance on a small sample of just nine Australian-listed firms is very likely to lead to substantial estimation error;

- The dispersion of risk estimates is reduced substantially as sample size is increased; and

- With a larger sample of firms, there is variation in risk estimates across samples and over time. Therefore, any regression-based estimate of risk should not be used in isolation to estimate the cost of capital.

3. Introduction

The Australian Energy Regulator ("AER" or "the regulator") is required to consider the use of relevant models to inform its cost of equity methodology. Three of these models are the Sharpe-Lintner Capital Asset Pricing Model, the Black Capital Asset Pricing Model ("CAPMs") and the Fama & French 3-factor model (Fama and French, 1993).

In implementing the CAPMs, particularly the Sharpe-Lintner CAPM, the regulator has considered estimates of the beta coefficient from ordinary least squares ("OLS") regression of stock returns on market returns (AER, 2009). The regulator has relied exclusively on a small sample of nine Australian-listed stocks in performing this regression. In this report, we examine the implications of sample size on the reliability of the historical risk estimates for the CAPMs and the Fama & French 3-factor model.

Essentially, the regulator faces a small sample problem. In estimating risk with reference to just nine comparable firms, there is the chance that those estimates would be very different if a larger sample of firms was available for analysis. What we can't tell from those nine stocks is what we might be missing because only a handful of comparable firms happened to be listed. In other words, we know the range of risk estimates for the small sample of nine stocks, and it is a wide range. But we don't really know what the risk estimate might have been if more stocks were available for analysis.

We use two approaches to measure the relationship between sample size and reliability.

The first approach is to measure the dispersion of risk estimates across different samples of firms from the same industry. The lower the variation in risk estimates across samples in the same industry, the more reliable the estimate. So we document the dispersion in average risk estimates across samples of firms from the same industry. We then demonstrate how this dispersion changes as more firms are introduced into each sample.

The second approach is to measure the variation in risk estimates over time for the same sample of firms. The greater the variation in regression-based risk estimates over time, the less reliance should be placed on the estimates. There are two reasons why risk estimates can vary over time. One reason is that firm risk is unchanged from one period to the next, but the risk estimate each period changes due to noise in the data. In this case of constant risk, clearly less noise is preferable. The other case is where the firm risk actually changes over time and the regression-based estimates actually capture these risk changes. But even in this case, the greater the variation over time in the risk estimate the less reliable will be that estimate. The reason for this is that the regression simply measures risk in the historical data. And even if that regression perfectly measured risk during the prior period, it is not suited to predict risk in the subsequent period.

Regardless of the reason why beta estimates change from one period to the next (noise or actual changes in risk) the more variation in this risk estimate the less reliance can be placed on the estimates for predicting risk in the future. So we measure the variation in risk estimates over time and report how this variation changes as sample size increases.

Stephen Gray and Jason Hall, SFG Consulting
Neil Diamond and Robert Brooks, Monash University

4. Method

4.1 Repeated sampling

We undertook a repeated random sampling exercise to examine the effect of such a small sample on the reliability of the estimates. We compiled a sample of 1,286 Australian-listed firms and split the sample into eight industry groups according to FTSE International Classification Benchmark (ICB) codes.[1] We also split our time period of 20 years into two periods of approximately 10 years and four periods of approximately five years.

We randomly selected a set of nine firms from each industry group and computed a beta estimate for each firm over the entire sample period, and for each of the 10-year and five-year sub-periods. We then compiled a mean beta estimate for all sub-samples. This mean estimate is analogous to the mean estimate across firms often referred to by regulators. Our objective is to demonstrate just how variable this estimate is (both within a time period and across multiple periods) when we are only able to observe nine firms, and how material the variability reduction is when we are able to observe a larger number of firms.

We repeat the selection of nine firms in each of the eight industry groups 1,000 times. This means that for each time period there are 72,000 beta estimates.[2] This results in 8,000 average industry beta estimates.[3] Note that in each sub-period of five or 10 years there are not necessarily nine firms in the sample, because not all firms are listed over the entire time period. So in some periods the sample will comprise just one firm and in other periods all nine firms will form part of the sample. This matches the small sample problem faced by the regulator, as not all nine comparable firms are available for analysis at any one time. We then repeat this analysis after selecting samples of 18, 27 and 36 firms in each industry.

We repeat this sampling exercise and compute coefficient estimates, and then risk premiums for the Fama & French model. In this instance we need to assess the reliability of all three coefficients jointly, but the magnitude of the coefficients themselves is not comparable. Specifically the Fama & French model says that the expected return is equal to the risk free rate, plus a premium for exposure to the market ($\beta \times MRP$), the size factor ($s \times SMB$) and the book-to-market factor ($h \times HML$). It is not appropriate to compare the variation of β, s, and h because the same change in each coefficient could have very different impacts on the cost of capital. For example, suppose *MRP* is estimated at 6% and *HML* is estimated at 2%. If the beta estimate is understated by 0.2, then the cost of equity capital will be understated by 1.2%. But if h is understated by 0.2, then the cost of equity capital will be understated by just 0.4%. So, rather than compare the reliability of the individual coefficients (β, s and h) we need to compare the reliability of the sum of the risk premiums.

4.2 Measures of reliability

As mentioned in the introduction we measure the reliability of risk estimates in two ways – the variation in beta estimates across samples in the same time period, and the variation in beta estimates over time for the same sample.

[1] There are two other industry groups (Telecommunications and Utilities) but there are not enough firms in these industry groups to perform meaningful analysis.

[2] That is, 9 sample firms per industry group × 8 industry groups × 1,000 simulations per industry group = 72,000 beta estimates.

[3] That is, 8 industry groups × 1,000 simulations per industry group = 8,000 average beta estimates.

4.2.1 Dispersion of risk estimates across samples in the same time period

We first measure the dispersion of average risk estimates in repeated samples of comparable firms from the same industry. The closer the average risk estimate of one sample of firms in an industry is to the average risk estimate of another small sample of firms in the same industry, the more reliable regression-based estimates of risk.

Furthermore, if increasing sample size improves the reliability of risk estimates then we should observe a reduction in this dispersion as sample size increases. So for each simulation we compute the mean risk estimate across sample firms, and compute the standard deviation of this mean risk estimate across the 1,000 simulations. This is the standard error of the risk estimate because it is the standard deviation of the mean estimates from repeated samples.

For example, suppose we selected two firms which had beta estimates of 0.6 and 1.2, respectively. This sample has an average beta estimate of 0.9. Then in the next simulation, suppose two firms in the same industry had beta estimates of 0.9 and 1.5, respectively. The second sample has an average beta estimate of 1.2. We repeat this analysis 1,000 times using nine firms and then compute the standard deviation of these average estimates. We then measure the standard deviation of mean estimates for sample sizes of 18, 27 and 36 firms.

4.2.2 Dispersion of risk estimates over time for the same sample

Our second measure of reliability relates to the dispersion of risk estimates over time. If regression-based risk estimates are suitable to use to predict future risk, then the average estimate for comparable firms should be reasonably stable from one period to the next. This time-series stability needs to be considered carefully in the context of cost of capital estimation using regression analysis of historical returns. The regression relies upon a measure of risk from historical data. That risk estimate is being used to estimate the cost of capital at a later point in time. Given this information, time-series stability implies greater reliability. If the true risk is stable, the more time-series variation in the risk estimate the less reliable will be that estimate. And even if regression analysis of stock returns was a perfect technique for measuring risk, the more the true risk varies over time the less reliable will be the measure of risk based upon historical data. The key point is that, in estimating cost of capital at a point in time, if regression-based risk estimates are to be used then the less time-series variation in the risk estimate the higher the reliability.

To measure the dispersion of risk estimates over time, we consider the four five-year sub-periods and the two 10-year sub-periods. Recall that for each simulation there is an average beta estimate for each industry group over each sub-period. In examining the four five-year sub-periods we compute the maximum average risk estimate from the four sub-periods, the minimum average risk estimate and take the difference. This tells us, on average, how much the risk estimate can vary from one five-year period to another. For example, suppose one particular sample had mean beta estimates of 0.7, 1.2, 0.8 and 0.9 over four five year periods. The highest beta estimate for this sample was 1.2 in the second period and the lowest beta estimate was 0.7 in the first period. So we report a difference between the maximum and minimum estimates of 0.5.

We repeat this analysis for the two 10-year sub-periods, in this case just computing the average risk estimates from one 10-year period to the next, and taking the difference. This tells us, on average, how much the risk estimate can vary from one 10-year period to another. For example, if a sample had a beta estimate of 1.3 from the first 10-year period and a beta estimate of 0.9 from the second 10-year period, the difference is 0.4.

4.3 Data

We computed stock returns on a four-weekly basis, but these were computed each week over the sample period. For example, in one sample a return from 1 January to 29 January would be one observation, and in another sample we could have a return from 8 January to 5 February. This means that the start points of each four-weekly return are random across samples. It also means that stock prices from each week in the sample period are used in computations, so we use a large amount of information.

To be eligible for inclusion in the dataset a stock needed to have at least 100 four-weekly returns available for analysis over the entire sample period. The market index is the All Ordinaries Index and we compute excess returns by subtracting the four-weekly equivalent of the yield to maturity on 10-year government bonds. Once we compute risk estimates we winsorize them at the 1st and 99th percentiles for each time period and industry. This means that, for observations below the 1st percentile and above the 99th percentile, we replace that observation with the 1st and 99th percentile. This mitigates the impact of some extreme risk estimates.[4]

In the table below we report descriptive statistics across the different sub-periods, along with the cut-off dates for each period. The reason these cut-off points do not correspond to exactly five years is because these time periods correspond to periods of available data for the nine Australian firms used by the AER. Of these nine firms, in the first five-year period, only one stock is available for analysis. In the subsequent three periods, there are between four and eight stocks available for analysis.

The table shows mean values for the beta estimates and the Fama & French estimates for each time period across the eight industries evaluated. These mean estimates are for the specific case where sample size is equal to 36, but the estimates for all samples are approximately the same because they are average values. In other words, because we are re-sampling from the same industry, the average beta estimate for the samples of nine firms will be the same as the average beta estimate for samples of 18, 27 and 36 firms. There are only minor deviations in the mean average beta estimates across different sample sizes in the same industry because we draw only 1,000 samples instead of an infinite number of samples.

If we consider the average beta estimates across the eight industries, these range from 0.65 for Consumer goods to 1.30 for Oil & gas, when estimated using up to 20 years of returns. The mean beta estimates vary substantially over time for each industry. For example, from the first 10-year period to the second 10-year period, the average beta estimate changes by 0.59 for Oil & Gas firms, 0.31 for Technology firms and 0.50 for Basic materials firms. Across all industries the range of changes in beta estimates is from 0.15 to 0.59.

[4] This serves to mitigate the impact of the most extreme risk estimates which would render the results meaningless if they were included. If the estimates were left unadjusted we run the risk of drawing conclusions from just 2% of the data rather than the remaining 98% of the data. Across the entire sample the maximum beta estimate was 250 and the minimum beta estimate was –119. Once we winsorize the data at the 1st and the 99th percentiles across time periods and industries, the maximum beta estimate is 22.6 and the minimum beta estimate is –20.0. Unadjusted, the standard deviation of beta estimates was 3.83 when estimated using up to five years of data, 0.95 when estimated using up to 10 years of data and 0.54 when estimated using up to 20 years of data. Adjusted, the corresponding standard deviations fall to 1.67, 0.91 and 0.53. With respect to the Fama & French risk premiums, the maximum estimate was 548% and the minimum estimate was –449%. Once we winsorize these estimate the maximum risk premium is 256% and the minimum estimate is –373%. Unadjusted, the standard deviation of risk estimates was 19.82% when estimated using up to five years of data, 11.72% when estimated using up to 10 years of data and 4.12% when estimated using up to 20 years of data. Adjusted, the corresponding standard deviations fall to 16.52%, 11.32% and 4.04%.

Table 1. Mean risk premium estimates

Industry	1st 5 yr period 1 May 92 to 28 Feb 97	2nd 5 yr period 1 Mar 97 to 31 Dec 01	3rd 5 yr period 1 Jan 02 to 31 Jan 07	4th 5 yr period 1 Feb 07 to 4 May 12	1st 10 yr period 1 May 92 to 31 Dec 01	2nd 20 yr period 1 Jan 02 to 4 May 12	20 year period 1 May 92 to 4 May 12
Panel A: Mean beta							
Oil & gas	1.85	0.68	1.14	1.46	0.77	1.36	1.30
Basic materials	1.02	0.80	1.12	1.45	0.85	1.35	1.24
Industrials	1.22	0.66	0.72	0.92	0.71	0.86	0.84
Consumer goods	1.02	0.50	0.71	0.81	0.57	0.76	0.65
Health care	1.98	1.17	1.09	1.22	1.21	0.97	1.01
Consumer services	1.09	0.88	0.80	0.80	0.93	0.78	0.82
Financials	-0.15	0.56	0.51	0.79	0.56	0.74	0.69
Technology	0.39	1.25	0.94	0.89	1.20	0.89	1.01
Panel B: Mean Fama & French risk premiums (%)							
Oil & gas	12.47	9.20	5.42	7.21	10.16	6.62	7.18
Basic materials	8.45	7.09	5.56	6.68	7.28	6.42	7.11
Industrials	9.81	5.18	5.98	5.54	5.65	5.91	5.87
Consumer goods	10.17	4.47	6.21	6.67	5.23	5.44	5.57
Health care	1.09	8.73	10.15	7.32	9.38	6.06	6.35
Consumer services	10.84	4.44	6.32	5.16	4.92	5.36	5.37
Financials	1.09	4.24	3.65	7.06	4.78	5.52	5.27
Technology	0.04	7.96	10.68	6.06	7.76	8.10	6.90
Panel C: Percentage of sample firms used in estimation (%)							
Oil & gas	56	81	98	95	81	99	100
Basic materials	65	82	97	89	82	97	100
Industrials	69	91	94	86	91	94	100
Consumer goods	71	94	97	79	94	97	100
Health care	43	76	98	95	78	99	100
Consumer services	63	89	97	83	89	97	100
Financials	65	86	95	85	87	95	100
Technology	36	96	100	89	96	100	100

With reference to the Fama & French model, recall that we refer to risk premiums in this case, rather than coefficients, because we need to assess the overall impact of three coefficients (β, s and h) and the magnitude of those coefficients can't be directly compared, and can't be simply added together. However, we can aggregate the risk premium associated with these coefficients, which is $\beta \times MRP + s \times SMB + h \times HML$.

We used S&P Australia value and growth indices in order to estimate *SMB* and *HML* over four-weekly periods. *SMB* is the return on the S&P Australia small stock index less the returns on the S&P Australia Medium and Large Capitalisation Index. *HML* is the return on the S&P Australia Value Index less the return on the S&P Australia Growth Index.[5] These are the factor returns we used to estimate the coefficients. We also used the annual average of these factor returns as estimates of the premiums associated with each factor, using the 23 years for which returns are available. These averages were 0.61% for *SMB* and 2.34% for *HML*. Hence, we compute the total risk premium as:

$$\text{Risk Premium} = \beta \times 0.0600 + s \times 0.0061 + h \times 0.0234$$

[5] This does not imply that these indices are the ideal indices to be used to measure exposure to the *SMB* and *HML* factors. These indices were merely selected as being readily-available over an extended time period. Our assumptions of 6.00% for *MRP*, 0.61% for *SMB* and 2.34% for *HML* also do not imply that these figures represent the best estimates for cost of capital estimation. The *MRP* assumption of 6.00% merely represents the current assumption of the AER, and the figures for *SMB* and *HML* are the long-term average figures in the indices used to perform the analysis. The point of our paper is not to provide estimates of risk exposure but to demonstate the likely variation in those estimates of risk exposure across different samples and time periods.

where beta is the estimated beta coefficient in the Fama & French model, *s* is the regression coefficient for *SMB* and *h* is the regression coefficient for *HML*.

Considering the Fama & French premiums, across industries the average premium ranges from 5.27% for Financials to 7.18% for Oil & gas firms, when estimated using up to 20 years of returns. Considering the changes in the Fama & French premiums from the first to the second 10-year period, the difference in the risk premium estimate from one period to the next ranges from 0.21% for Consumer goods to 3.54% for Oil & gas.

In the last section of the table we document the average proportion of sample firms which were available for analysis in each sub-period. For example, if the sample size is nine and there are six firms available for analysis in a particular time period, then 67% of sample firms are available for analysis is that time period. In constructing our repeated samples we have attempted to replicate the challenge faced by the regulator, which not only has a small sample of firms available for analysis, but at no stage are all of those firms actually listed at the same time. So early in the sample period there are considerably fewer than 9, 18, 27 or 36 firms in each sample, because a large number of firms are not listed until later.

5. Results

5.1 CAPM beta estimates

5.1.1 Dispersion of average CAPM beta estimates

In Table 2 we document the dispersion of average beta estimates as the number of sample firms increases. Each cell in the table represents the standard deviation of the average beta estimate across 1,000 samples of firms drawn from the same industry. This can be termed the standard error of the beta estimate, which we discuss further in a subsequent section. The standard error is the standard deviation of a mean estimate. What the table demonstrates is that by increasing the sample size from nine firms to 18 firms there is approximately a 30% reduction in the dispersion of the average beta estimates. This is the average percentage reduction in the standard error as we move from the 9 firm column to the 18 firm column. The reduction in estimation error is the same regardless of whether the beta estimates are computing using up to five, 10 or 20 years of returns. Specifically, the average reductions in standard errors are 33%, 31% and 32% over these three estimation windows.

As sample size increases to 27 firms there is a 47% reduction in the standard error. At a sample size of 36 firms, there is a reduction of 57. These are the average percentage changes over estimation of five, 10 and 20 years. In short, by increasing the sample size from nine to 27 firms the dispersion of average beta estimates is cut in half.

To place this in context, consider the figures reported in Panel F, which corresponds to the most recent 10-year period from 1 January 2002 to 4 May 2012. Across the eight industries, the standard error of beta estimates from a sample of nine firms ranges from 0.15 to 0.22. This is a measurement of how variable the average beta estimates could be if we just happened to observe a different set of nine firms in that industry. It can also be thought of as a measurement of the difference in average beta estimate we could observe if we just happened to observe a different random set of company-specific events for the same nine firms. This means that a range of outcomes one standard error either side of the mean could have a magnitude of somewhere from 0.30 to 0.44 on the beta estimate, depending upon industry.

However, if we are able to observe 18 firms there is less chance that the average beta estimate will be adversely affected by company-specific events. In the next column we see the standard error of beta estimates across industries lies within a range of 0.11 to 0.15. So a range for beta estimates which is one standard error either side of the mean is 0.22 to 0.30, depending upon industry.

This analysis relates to how point estimates and confidence intervals in small samples should be interpreted. In its analysis of 2009, the AER discussed confidence intervals, but it was unclear just how those confidence intervals led to its two critical conclusions, namely:

- Stock return data suggested a range for the beta estimate, re-geared to 60%, of 0.41 to 0.68. The lower and upper bound of that range were two average values from different studies.
- A beta estimate of 0.80 would be adopted, because this would encourage investment, having regard to the potential consequences of underinvestment, and to promote regulatory stability.

It is important to recognize that previous submissions in relation to confidence intervals and the AER's interpretation of those submissions were framed with reference to the NER requirement to present persuasive evidence to depart from a prior view. The persuasive evidence test has been removed from the NER. So we consider the issue of confidence intervals from first principles.

Table 2. Standard error of CAPM beta estimates according to different sample sizes

Panel A: First five-year period (1 May 1992 to 28 February 1997)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	1.19	0.87	0.64	0.49
Basic materials	1.49	0.99	0.76	0.65
Industrials	1.39	0.67	0.49	0.44
Consumer goods	0.99	0.67	0.49	0.38
Health care	2.38	1.42	1.08	0.84
Consumer services	1.30	0.77	0.64	0.49
Financials	1.51	1.10	0.88	0.76
Technology	1.94	1.15	0.85	0.67

Panel B: Second five-year period (1 March 1997 to 31 December 2001)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.48	0.33	0.24	0.20
Basic materials	0.34	0.24	0.19	0.16
Industrials	0.37	0.25	0.20	0.17
Consumer goods	0.22	0.15	0.12	0.10
Health care	0.62	0.40	0.33	0.26
Consumer services	0.35	0.23	0.18	0.15
Financials	0.31	0.22	0.18	0.15
Technology	0.44	0.29	0.21	0.17

Panel C: Third five-year period (1 January 2002 to 31 January 2007)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.34	0.23	0.17	0.14
Basic materials	0.34	0.25	0.20	0.17
Industrials	0.26	0.18	0.15	0.12
Consumer goods	0.28	0.19	0.15	0.11
Health care	0.35	0.23	0.18	0.14
Consumer services	0.28	0.19	0.16	0.12
Financials	0.26	0.18	0.14	0.11
Technology	0.26	0.17	0.12	0.10

Panel D: Fourth five-year period (1 February 2007 to 4 May 2012)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.24	0.17	0.13	0.11
Basic materials	0.23	0.16	0.14	0.11
Industrials	0.20	0.15	0.11	0.10
Consumer goods	0.32	0.22	0.17	0.14
Health care	0.74	0.51	0.39	0.31
Consumer services	0.30	0.21	0.17	0.14
Financials	0.25	0.17	0.14	0.14
Technology	0.17	0.12	0.08	0.06

Panel E: First 10-year period (1 May 1992 to 31 December 2001)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.50	0.34	0.25	0.20
Basic materials	0.32	0.23	0.18	0.16
Industrials	0.36	0.23	0.19	0.16
Consumer goods	0.19	0.13	0.10	0.08
Health care	0.61	0.39	0.33	0.26
Consumer services	0.33	0.22	0.17	0.14
Financials	0.27	0.21	0.16	0.14
Technology	0.43	0.28	0.21	0.16

Panel F: Second 10-year period (1 January 2002 to 4 May 2012)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.22	0.14	0.11	0.09
Basic materials	0.21	0.15	0.12	0.10
Industrials	0.18	0.12	0.09	0.08
Consumer goods	0.22	0.15	0.12	0.09
Health care	0.17	0.11	0.09	0.07
Consumer services	0.18	0.13	0.12	0.08
Financials	0.19	0.13	0.11	0.09
Technology	0.15	0.11	0.07	0.06

Panel G: Entire 20-year (1 May 1992 to 4 May 2012)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.15	0.09	0.07	0.06
Basic materials	0.17	0.12	0.10	0.08
Industrials	0.15	0.11	0.08	0.07
Consumer goods	0.13	0.09	0.07	0.06
Health care	0.17	0.11	0.09	0.07
Consumer services	0.15	0.11	0.09	0.07
Financials	0.15	0.10	0.08	0.07
Technology	0.16	0.11	0.08	0.06

The confidence interval around a mean is the range of average outcomes we would expect to observe if the analysis was repeated many times over. In most cases we only have one sample, not many samples, so we use equations and assumptions to estimate this range of mean outcomes.[6] For example, in a sample of nine firms, if the mean beta estimate was 1 and the standard deviation was 0.6, the standard error is 0.2 (that is, $0.6 \div \sqrt{9} = 0.6 \div 3 = 0.2$). The standard error is an estimate of the standard deviation of averages we would observe if an experiment was repeated many times. When the number of observations is nine, a 90% confidence interval is 1.86 standard errors either side of the mean, so the 90% confidence interval is 0.63 to 1.37.

The problem is that, in small samples, the actual confidence interval from repeated sampling can be much wider than this range. What the repeated sampling analysis does is allow us to estimate the standard error directly from the data, rather than being computed from one sample. Recall that the standard error is the standard deviation of mean values from repeated experiments. That is what is presented in Table 2 – the standard deviation of average values if computed many times. What it shows is that the typical standard error in beta estimates for samples of nine firms is around 0.15 to 0.22, when beta estimates are computed over the most recent 10-year period. This standard error can be effectively cut in half if the sample size is increased to 27 or more firms.

Furthermore, given that we create our own distribution of sample means by repeated sampling we do not need to rely upon an equation to determine a confidence interval. We can report the 90% confidence interval as the range from the 5th to the 95th percentile of the 1,000 sample means computed from directly from the data. In the table below, we present these confidence intervals for the second 10-year period, which corresponds to Table 4, Panel F.

[6] For example, in estimating a confidence interval using a t-distribution we assume that the observations in the sample are independent of one another and that the population of data from which the sample is drawn is normally distributed.

Table 3. Confidence interval from repeated samples

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.91 to 1.66	1.09 to 1.57	1.16 to 1.52	1.21 to 1.50
Basic materials	0.99 to 1.61	1.11 to 1.58	1.16 to 1.55	1.19 to 1.51
Industrials	0.55 to 1.13	0.65 to 1.05	0.71 to 1.01	0.72 to 0.98
Consumer goods	0.44 to 1.17	0.54 to 1.04	0.58 to 0.97	0.62 to 0.92
Health care	0.68 to 1.26	0.79 to 1.16	0.83 to 1.13	0.86 to 1.09
Consumer services	0.49 to 1.10	0.58 to 1.01	0.55 to 0.95	0.64 to 0.92
Financials	0.43 to 1.05	0.51 to 0.96	0.57 to 0.92	0.59 to 0.89
Technology	0.66 to 1.15	0.72 to 1.07	0.77 to 1.01	0.80 to 0.99

Time period is from 1 January 2002 to 4 May 2012. Confidence interval is from the 5th percentile to the 95th percentile of industry mean beta estimates. The confidence intervals correspond to the standard errors reported in Table 4, Panel F, but are computed directly from the data.

What the table demonstrates is that, if we rely upon just nine sample firms for analysis, the confidence intervals take on a very wide range, from 0.50 for Technology firms to 0.75 for Oil & gas firms. The width of these confidence intervals is around 1.65 standard errors (the figures reported in Table 2, Panel F.) All of the confidence intervals encompass the value of one, but there is a 5% chance of observing a sample beta estimate which is less than 0.43 (Financials) and a 5% chance of observing a sample beta estimate which is more than 1.66 (Oil & gas).

If we increase the sample size to 18 firms the width of the confidence interval narrows to around 0.35 to 0.49, and it falls further again to a width of around 0.24 to 0.39 if sample size is increased to 27 firms. It is only once we increase sample size to 36 firms that we start to observe most confidence intervals lying entirely above or below one. The implication of these confidence intervals is that if beta estimates are made from regression analysis of just nine sample firms, there is a high probability this estimate is very different from the risk in the industry. There is a high chance that reliance on this point estimate in isolation will lead to a material mis-statement of the cost of capital.

5.1.2 Variation in CAPM beta estimates over time

In the previous section we documented how the dispersion of mean beta estimates significantly decreases as more firms are incorporated into the analysis. In this section we summarise the variation in beta estimates over time, and the extent of how that variation is reduced as sample size increases.

We have beta estimates computed over periods of five, 10 and 20 years. If regression-based beta estimates are unstable over time, and this accurately reflects true changes in risk, then it is clear that a beta estimate from historical returns would be a poor predictor of future systematic risk. Alternatively, if the true level of systematic risk is stable, but the beta estimate fluctuates due to noise in returns, then this is merely a noisy estimate of systematic risk.

To measure the variation in beta estimates over time when estimated using up to five years of returns, we computed the highest and lowest average industry beta estimates for each simulation over the four five-year periods. We then take the difference between these average estimates. In Table 4, Panel A, we present the average of these differences in high and low estimates. We repeat the exercise using beta estimates computed using 10 years of returns and these estimates are presented in Panel B.

When just nine firms are used in estimation, the average difference in high versus low mean beta estimates across industries over five-year periods lies within the range of 1.03 to 2.02. Consider the value of 2.02 for Health care. This means that, if we look at the four average beta estimates over different five year periods for a sample of nine firms, the average difference between the highest versus the lowest average estimate is 2.02. The average high estimate was 2.58 and the average low estimate was 0.55.

Table 4. Variation in average CAPM beta estimates over time

Panel A: Average difference between highest and lowest mean beta estimates over four five-year periods

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	1.60	1.43	1.28	1.23
Basic materials	1.53	1.17	0.99	0.92
Industrials	1.27	0.86	0.71	0.70
Consumer goods	1.03	0.84	0.70	0.64
Health care	2.02	1.56	1.37	1.16
Consumer services	1.17	0.82	0.73	0.58
Financials	1.35	1.23	1.09	1.05
Technology	1.75	1.28	1.07	0.98

Panel B: Average difference between mean beta estimates over two 10-year periods

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	0.62	0.61	0.58	0.59
Basic materials	0.53	0.49	0.49	0.50
Industrials	0.31	0.24	0.20	0.18
Consumer goods	0.24	0.21	0.20	0.20
Health care	0.50	0.38	0.33	0.29
Consumer services	0.29	0.22	0.21	0.17
Financials	0.29	0.26	0.21	0.21
Technology	0.43	0.35	0.34	0.31

As we increase sample size this range of outcome decreases. If 18 firms are included in the sample the average difference in high versus low mean beta estimates lies within the range of 0.82 to 1.56. This represents a reduction in dispersion of 9% to 32%. Further increasing the sample to 27 firms reduces the range to 0.70 to 1.37, which is a reduction in dispersion of 19% to 44%.

Despite this improvement from including additional firms, there is still considerable time-series variation in beta estimates. The key point for cost of capital estimation is that, depending upon which five-year period of stock returns we happen to observe, the estimate of beta can vary markedly. During one five-year period an industry could be considered to have systematic risk well below the market average of one. During another five-year period within the same two decades, that industry could be considered to have systematic risk well above the market average of one.

If we extend the estimation period to 10 years and compute the difference in average beta estimates, we observe an average difference across industries within the range of 0.24 to 0.62. So despite having twice the returns observations to use, on average an industry's beta estimate from one decade to the next would be expected to change by around 0.24 to 0.62. Extending the number of sample firms to 27 reduced the range to 0.20 to 0.58, which represented a reduction in dispersion of 2% to 24%. At 36 firms, the range is 0.17 to 0.59.

This still represents substantial variation in beta estimates over time, especially if we consider that this represents an average result. This means that in the average case, even when a sample of 36 firms is used and the estimates are computed using 10 years of returns, the industry beta estimate will change by 0.17 to 0.59 from one decade to the next.

For clarity, we repeat our point that this time-series variation has adverse consequences for cost of capital estimation even if the regression analysis perfectly measured risk in each time period. For example, suppose that, in reality, the systematic risk of 27 Consumer Goods firms changed by 0.20 from one 10-year period to the next. That is, suppose that the actual beta rose from 0.57 to 0.77. If we were to use regression analysis to estimate beta, we would use the figure of 0.57 to compute the cost of capital when in reality next period the beta is 0.77. We would understate beta by 0.20, and if applied to a market risk premium of 6.00%, would understate the cost of equity by 1.20%. So regardless of whether the beta estimates change because of noise in the data, or because of true changes in risk, stability in the regression-based beta estimates is desirable for estimating the cost of capital.

5.1.3 Variation in CAPM beta estimates over time for lowest dispersion estimates

We performed additional analysis to determine whether our results were driven by our selection from broad industry groups. Using beta estimates from the first 10-year period, we partitioned the sample into deciles within each industry according to the standard deviation of beta estimates within each sub-sample. So decile one contains the 100 cases within each industry in which the beta estimates were most tightly clustered together. If regression-based beta estimates are reliable for cost of capital estimation, then we should observe the most stability over time for these cases in which the estimates are closest together. We separately analysed the first three deciles formed on this basis.

In Table 5 we present the mean beta estimates over the first and second 10-year periods, and the average of the absolute difference between these estimates. So the figures under the "Diff" column correspond to the figures reported in Table 4, Panel B, for the whole sample.[7] We observe that even amongst the cases with beta estimates closest to each other, there is substantial variation in beta estimates from one decade to the next. In the first decile, when samples are formed from nine firms, the average absolute difference in mean beta estimates from one period to the next ranges from 0.18 for Consumer services to 0.50 for Oil & gas firms. If we consider the second and third deciles, the corresponding ranges are 0.16 to 0.51 and 0.19 to 0.51.

If we expand the sample size we observe less variation in beta estimates over time, but this variation is still material. When 18 firms are used in the analysis, amongst the first decile the beta estimates vary from 0.13 for Consumer services to 0.45 for Technology. This represents a reduction in variation over time of 6% to 27%.[8] Expanding the sample to 27 firms the range is 0.11 to 0.48 for this decile, which represents a reduction in variation over time of 2% to 40%. If we consider the second and third deciles, there is a similar reduction in variation over time.

The implication of these results is that even if a set of comparable firms is compiled from within the same industry, and even if those firms have similar beta estimates compiled from a prior period, on average there will be a substantial difference in the beta estimates for that sample if the analysis was performed using another ten years of data. When that sample uses just nine firms in the analysis, this difference is likely to be around 0.34, which is the average difference in beta estimates across the first three deciles and across all eight industries. When the sample relies upon 18 firms or 27 firms, the difference is likely to be around 0.30 or 0.28, respectively.

[7] The average absolute difference does not equal the difference between the mean beta estimate from period one and period two, because it is the average of the absolute difference between the beta estimates.

[8] That is, the average absolute difference in mean beta estimates for Basic Materials firms fell from 0.49 to 0.46, a reduction of 6%. For Consumer services the corresponding variation over time fell from 0.18 to 0.13, a reduction of 27%.

Table 5. Variation in CAPM beta estimates over time for low dispersion cases

Panel A: First decile according to standard deviation of beta estimates in the first 10-year period

	9 firms			18 firms			27 firms			36 firms		
Industry	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|
Oil & gas	0.94	1.40	0.50	0.95	1.37	0.43	0.99	1.35	0.36	0.96	1.38	0.42
Basic materials	0.84	1.32	0.49	0.87	1.33	0.46	0.90	1.36	0.48	0.89	1.37	0.48
Industrials	0.62	0.83	0.24	0.60	0.79	0.21	0.63	0.85	0.22	0.65	0.82	0.18
Consumer goods	0.51	0.73	0.26	0.54	0.73	0.20	0.57	0.77	0.21	0.58	0.77	0.20
Health care	1.24	0.98	0.34	1.27	0.99	0.31	1.28	0.99	0.30	1.31	0.99	0.32
Consumer services	0.72	0.78	0.18	0.72	0.78	0.13	0.76	0.75	0.11	0.78	0.79	0.09
Financials	0.55	0.75	0.25	0.57	0.74	0.19	0.60	0.75	0.17	0.58	0.73	0.17
Technology	1.37	0.90	0.49	1.35	0.90	0.45	1.35	0.90	0.45	1.28	0.90	0.39

Panel B: Second decile according to standard deviation of beta estimates in the first 10-year period

	9 firms			18 firms			27 firms			36 firms		
Industry	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|
Oil & gas	0.93	1.35	0.47	1.00	1.35	0.37	0.97	1.36	0.39	0.95	1.38	0.43
Basic materials	0.87	1.35	0.49	0.88	1.36	0.48	0.90	1.37	0.47	0.90	1.34	0.45
Industrials	0.60	0.80	0.25	0.61	0.85	0.26	0.71	0.85	0.15	0.72	0.85	0.14
Consumer goods	0.56	0.71	0.21	0.57	0.78	0.22	0.59	0.76	0.17	0.58	0.77	0.19
Health care	1.27	0.95	0.38	1.30	1.00	0.31	1.36	0.99	0.37	1.37	0.98	0.39
Consumer services	0.73	0.76	0.16	0.77	0.78	0.13	0.79	0.76	0.10	0.83	0.79	0.09
Financials	0.58	0.74	0.24	0.62	0.76	0.19	0.60	0.75	0.18	0.58	0.74	0.19
Technology	1.39	0.92	0.51	1.35	0.89	0.46	1.30	0.90	0.40	1.22	0.89	0.33

Panel B: Third decile according to standard deviation of beta estimates in the first 10-year period

	9 firms			18 firms			27 firms			36 firms		
Industry	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|
Oil & gas	0.95	1.35	0.44	0.95	1.36	0.42	0.97	1.37	0.40	0.97	1.37	0.40
Basic materials	0.87	1.37	0.52	0.91	1.36	0.45	0.93	1.35	0.43	0.89	1.35	0.46
Industrials	0.62	0.78	0.24	0.69	0.81	0.19	0.76	0.85	0.14	0.72	0.85	0.15
Consumer goods	0.57	0.76	0.22	0.60	0.76	0.17	0.59	0.79	0.20	0.59	0.79	0.20
Health care	1.30	0.97	0.39	1.38	0.98	0.41	1.42	0.99	0.43	1.34	0.98	0.36
Consumer services	0.80	0.81	0.19	0.83	0.77	0.14	0.87	0.78	0.13	0.87	0.78	0.12
Financials	0.63	0.74	0.19	0.62	0.74	0.20	0.58	0.74	0.18	0.57	0.75	0.20
Technology	1.33	0.90	0.51	1.27	0.90	0.39	1.27	0.89	0.39	1.24	0.89	0.35

5.2 Fama & French risk premium estimates

5.2.1 Dispersion of average risk premium estimates

In Table 6 we document the dispersion of Fama & French average risk premium estimates as the number of sample firms increases. Each cell in the table represents the standard error of the risk premium estimate across 1,000 simulations of firms drawn from the same industry. As we increase sample size we observe approximately the same reduction in dispersion of estimates as we observed with beta estimation. As sample size increases from nine to 18 firms, the standard error is generally reduced by one-third, and as sample size increases to 27 firms the standard error is reduced to approximately half.

Table 6. Standard error of risk premium estimates according to different sample sizes (%)

Panel A: First five-year period (1 May 1992 to 28 February 1997)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	9.60	6.45	4.84	3.85
Basic materials	8.86	5.98	4.83	4.08
Industrials	8.30	5.68	4.58	3.67
Consumer goods	9.90	6.21	4.71	3.77
Health care	37.35	20.97	16.43	12.85
Consumer services	12.49	8.71	6.51	5.46
Financials	10.19	7.14	5.51	4.65
Technology[9]	17.68	11.19	8.32	6.58

Panel B: Second five-year period (1 March 1997 to 31 December 2001)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	6.92	4.73	3.64	2.90
Basic materials	4.49	3.22	2.47	2.11
Industrials	3.55	2.40	1.98	1.61
Consumer goods	3.36	2.26	1.70	1.38
Health care	7.80	5.11	3.95	3.04
Consumer services	3.17	2.21	1.59	1.35
Financials	3.34	2.17	1.79	1.50
Technology	6.14	3.95	3.00	2.32

Panel C: Third five-year period (1 January 2002 to 31 January 2007)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	3.24	2.09	1.64	1.29
Basic materials	3.71	2.57	1.99	1.78
Industrials	2.73	1.99	1.54	1.26
Consumer goods	2.93	1.96	1.55	1.19
Health care	2.81	1.87	1.46	1.21
Consumer services	2.54	1.78	1.30	1.09
Financials	2.36	1.67	1.33	1.16
Technology	5.03	3.62	2.71	2.07

Panel D: Fourth five-year period (1 February 2007 to 4 May 2012)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	1.77	1.13	0.89	0.73
Basic materials	2.11	1.49	1.19	0.98
Industrials	2.32	1.50	1.21	1.02
Consumer goods	2.11	1.34	1.07	0.84
Health care	2.89	3.47	2.65	2.20
Consumer services	3.07	2.01	1.56	1.25
Financials	2.99	2.16	1.54	1.26
Technology	2.92	1.86	1.39	1.05

[9] At the time of writing we were missing data for this industry sector when nine firms are used in the analysis. This will be included in the final report.

Panel E: First 10-year period (1 May 1992 to 31 December 2001)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	6.58	4.50	3.49	2.75
Basic materials	4.36	3.12	2.43	2.06
Industrials	3.34	2.21	1.88	1.50
Consumer goods	3.23	2.16	1.63	1.31
Health care	7.75	5.07	3.90	3.04
Consumer services	3.08	2.14	1.55	1.32
Financials	2.98	1.95	1.61	1.36
Technology	6.07	3.88	2.95	2.28

Panel F: Second 10-year period (1 January 2002 to 4 May 2012)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	1.37	0.87	0.69	0.57
Basic materials	1.84	1.32	1.05	0.89
Industrials	1.60	1.09	0.87	0.74
Consumer goods	2.48	1.55	1.26	1.02
Health care	1.79	1.24	0.94	0.74
Consumer services	2.05	1.32	1.04	0.85
Financials	1.85	1.28	0.98	0.86
Technology	4.66	3.43	2.55	1.95

Panel G: Entire 20-year (1 May 1992 to 4 May 2012)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	1.05	0.72	0.56	0.46
Basic materials	1.35	0.95	0.74	0.65
Industrials	1.17	0.81	0.63	0.52
Consumer goods	1.15	0.73	0.59	0.48
Health care	1.56	1.05	0.82	0.66
Consumer services	1.26	0.85	0.66	0.54
Financials	1.16	0.81	0.63	0.57
Technology	1.45	0.94	0.68	0.54

To illustrate this more clearly we again refer to the ten-year period from 1 January 2002 to 4 May 2012, presented in Panel F. When nine firms are used to estimate the Fama & French risk premium, the standard error across industries, excluding Technology, ranges from 1.37% to 2.48%. The standard errors for Technology are materially greater than the standard errors for the other industries but the overall result is the same. This industry sector exhibits the same reduction in dispersion of outcome as the other industries when sample size increases. The implication is that by relying on such a small sample of firms, the cost of capital estimate could easily vary from ±2% from one sample to the next, and this is just one standard error.

If we expand sample size to 18 firms there is much less chance that the risk premium estimates will be adversely affected by random events that are not associated with exposure to risk factors. The standard error of risk premium estimates, excluding Technology, lies within the range of 0.87% to 1.55%. While there is still considerable dispersion in the risk premium estimates from different samples, this is a substantial reduction in the dispersion of outcomes from relying upon just nine firms. Further increasing sample size to 27 firms results in most industries having a standard error of risk premiums of around 1%.

Our earlier discussion with respect to confidence intervals remains relevant here. What the figures in Table 6 represent is an estimate of the standard error of risk premiums across industries and for different sample sizes. The standard error can be effectively reduced by one-third by increasing sample size from nine to 18 firms, and effectively reduced by half by increasing sample size to 27 firms.

Table 7. Risk premium confidence interval from repeated samples (%)

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	4.37 to 8.80	5.14 to 8.03	5.52 to 7.83	5.64 to 7.54
Basic materials	3.39 to 9.46	4.29 to 8.52	4.74 to 8.19	4.92 to 7.79
Industrials	3.20 to 8.52	4.08 to 7.69	4.54 to 7.39	4.70 to 7.13
Consumer goods	0.55 to 8.91	2.63 to 7.77	3.32 to 7.44	3.72 to 7.04
Health care	3.24 to 9.29	4.12 to 8.14	4.62 to 7.70	4.84 to 7.29
Consumer services	1.79 to 8.64	3.10 to 7.37	3.66 to 7.13	3.97 to 6.79
Financials	2.34 to 8.50	3.23 to 7.41	3.92 to 7.14	4.00 to 6.93
Technology	4.02 to 21.00	4.84 to 14.55	5.25 to 12.17	5.59 to 10.72

Time period is from 1 January 2002 to 4 May 2012. Confidence interval is from the 5th percentile to the 95th percentile of industry mean risk premium estimates. The confidence intervals correspond to the standard errors reported in Table 6, Panel F, but are computed directly from the data.

In Table 7 we report 90% confidence intervals pertaining to the most recent 10 year period. We see that with samples of just nine firms the estimate of the risk premium takes on a wide range. Excluding Technology firms the width of the confidence interval lies from 4.43% to 8.36%, and on average is about 6%. By incorporating 18 firms in estimation the width of the confidence intervals falls to about 4% on average, and falls again to about 3% if 27 firms are incorporated in the analysis.

5.2.2 Variation in risk premium estimates over time

As with beta estimation, a useful property of the estimates of the Fama & French coefficients is that they are reasonably stable over time. If the estimates vary substantially from one period to the next because the true risk exposure changes, then regression-based analysis will not be a reliable predictor of future risk exposure. And if the estimates change materially from period to period because of noise in the data, this suggests that noise may outweigh the estimates' usefulness.

In Table 8, Panel A, we report the average difference between high and low sample averages over the four five-year periods for each industry. For six of the eight industries, when just nine firms are used in estimating the coefficients, this difference is within the range of 9.67% to 12.21%. The difference is considerably larger for Health care and Technology stocks but the impact of increasing sample size is the same. This means that, if the Fama & French risk premiums were estimated four times over 20 years in five year periods, the highest estimate would be around 10% to 12% different to the lowest estimate.

As we increase sample size to 18 firms, this difference between high and low estimates is reduced. For six of the eight industries the difference in risk premium estimates over time is now within the range of 7.63% to 9.65%, or approximately 8% to 10%. Incorporating 27 firms into the analysis sees the range fall to 6.23% to 8.42%, or approximately 6% to 8%.

What this means in general terms is that, selecting from a set of nine firms, the estimate of the risk premium from one five year period could be about 11% different to the estimate of the risk premium in another period, if the analysis was repeated three more times. This difference falls to about 9% if 18 firms are used in the analysis and about 7% if 27 firms are used in the analysis.

If we extend the estimation period to 10 years and compute the difference in average risk premium estimates, we observe an average difference across these six industries within the range of 2.87% to 4.58%, or approximately 3% to 5%, when nine firms are used in analysis. This range falls to 1.97% to 4.06% when sample size increases to 18 firms, and then to 1.53% to 3.81% when 27 firms are used. So the variation in risk premium estimates from one 10-year period to the next is about 4% when nine firms are analysed, about 3% when 18 firms are analysed and less than 3% when 27 firms are analysed.

Stephen Gray and Jason Hall, SFG Consulting
Neil Diamond and Robert Brooks, Monash University

Table 8. Variation in risk premium estimates over time

Panel A: Average difference between highest and lowest mean risk premium estimates over four five-year periods

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	12.21	9.65	8.42	7.89
Basic materials	10.76	7.82	6.23	5.56
Industrials	9.67	7.63	6.53	5.94
Consumer goods	9.89	7.93	7.19	6.27
Health care	23.19	17.54	15.81	13.66
Consumer services	9.70	8.47	7.35	7.30
Financials	9.83	8.29	7.40	7.02
Technology	16.36	13.79	12.58	11.62

Panel B: Average difference between mean risk premium estimates over two 10-year periods

Industry	9 firms	18 firms	27 firms	36 firms
Oil & gas	4.58	4.06	3.81	3.64
Basic materials	3.66	2.87	2.23	1.96
Industrials	2.87	1.97	1.70	1.40
Consumer goods	3.09	2.11	1.64	1.33
Health care	6.53	4.87	4.15	3.76
Consumer services	2.90	1.98	1.53	1.32
Financials	2.95	2.05	1.65	1.46
Technology	5.94	4.04	3.12	2.42

5.2.3 Variation in risk premium estimates over time for lowest dispersion estimates

As with our analysis of CAPM beta estimates we performed additional analysis to determine whether our results were driven by our selection from broad industry groups. Using risk premium estimates from the first 10-year period, we partitioned the sample into deciles within each industry according to the standard deviation of risk premium estimates within each sub-sample. So decile one contains the 100 cases within each industry in which the risk estimates were most tightly clustered together. If regression-based estimates of risk are reliable for cost of capital estimation, then we should observe the most stability over time for these cases in which the estimates are closest together. We separately analysed the first three deciles formed on this basis. We present results in Table 9. The figures in the columns labelled "Diff" correspond to those reported in Table 8, Panel B, for the full sample.

Considering first the decile one results, and a sample of just nine firms, the difference in the estimated risk premium from one period to the next ranges from 1.89% for Industrials to 2.74% for Oil & gas firms (excluding Health care and Technology firms). This means that, even for the 10% of cases in which the initial risk premium estimate was closest together, the risk premium estimate in the next 10-year period was different by around 1.89% to 2.74% across industries. There is a similar degree of variation over time for stocks in deciles two and three, with corresponding ranges of 1.87% to 2.47%, and 1.76% to 2.56%, respectively.

If we increase sample size to 18 firms, the variation in risk premium estimates over time falls to a range of 1.44% to 2.43%. At 27 firms the corresponding range is lower again, at 1.25% to 2.13%. This deviation is still material in the context of cost of capital estimation. The implication is that even if we have a reasonably large set of comparable firms from the same industry, with initial risk premium estimates that are close in magnitude, on average the estimated risk premium would be different by around 1.25% to 2.13% purely because a different ten years of data was analysed.

Table 9. Variation in risk premium estimates over time for cases in which risk premium estimates have low dispersion

Panel A: First decile according to standard deviation of beta estimates in the first 10-year period

	9 firms			18 firms			27 firms			36 firms		
Industry	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|
Oil & gas	8.08	6.54	2.74	7.99	6.79	1.63	7.79	6.79	1.31	7.98	6.77	1.35
Basic materials	8.66	6.58	2.69	8.44	6.37	2.43	8.39	6.42	2.13	8.59	6.42	2.31
Industrials	4.92	6.02	1.89	5.11	6.06	1.44	5.41	6.20	1.36	5.51	6.01	1.05
Consumer goods	3.80	4.79	2.72	4.31	5.30	1.80	4.55	5.21	1.31	4.91	5.51	1.08
Health care	7.97	5.58	3.25	9.51	5.98	3.64	9.48	6.42	3.15	10.09	6.32	3.79
Consumer services	4.14	5.71	2.09	4.64	5.33	1.67	5.00	5.39	1.25	5.31	5.42	0.83
Financials	4.12	5.39	2.08	4.24	5.20	1.53	4.43	5.42	1.35	4.43	5.54	1.34
Technology	10.53	8.20	4.93	11.03	8.37	4.20	10.90	8.46	3.21	10.81	8.21	2.82

Panel B: Second decile according to standard deviation of beta estimates in the first 10-year period

	9 firms			18 firms			27 firms			36 firms		
Industry	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|
Oil & gas	8.26	6.73	2.34	8.02	6.83	1.73	7.95	6.68	1.44	8.53	6.76	1.80
Basic materials	8.16	6.53	2.47	8.52	6.40	2.73	8.39	6.35	2.31	8.08	6.36	1.89
Industrials	5.22	5.81	1.87	5.26	6.06	1.30	5.79	5.97	1.26	5.67	5.90	1.22
Consumer goods	4.16	5.30	2.24	4.79	5.20	1.38	5.03	5.35	1.15	5.00	5.46	1.03
Health care	8.83	5.97	3.35	9.74	6.21	3.90	9.49	6.33	3.37	9.91	6.20	3.77
Consumer services	4.53	5.14	2.30	5.44	5.35	1.25	5.48	5.25	1.31	5.31	5.24	1.22
Financials	4.84	5.20	2.15	4.34	5.17	1.54	4.42	5.51	1.54	4.28	5.45	1.56
Technology	11.61	7.42	5.55	10.59	8.07	3.95	10.44	8.11	3.24	10.02	8.07	2.41

Panel B: Third decile according to standard deviation of beta estimates in the first 10-year period

	9 firms			18 firms			27 firms			36 firms		
Industry	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|	Mn 1	Mn 2	|Diff|
Oil & gas	7.63	6.89	1.95	7.97	6.86	1.43	8.58	6.74	1.92	8.76	6.68	2.15
Basic materials	8.42	6.68	2.56	8.34	6.46	2.48	8.16	6.57	1.80	8.35	6.38	2.24
Industrials	5.13	6.10	1.76	5.79	6.02	1.49	5.54	5.84	1.71	5.64	5.93	1.26
Consumer goods	4.94	5.22	2.20	5.26	5.47	1.45	5.26	5.54	1.34	4.27	5.57	1.53
Health care	9.86	6.37	3.88	9.78	6.48	3.70	9.73	6.29	3.72	10.00	6.18	3.84
Consumer services	4.85	4.84	2.04	5.66	5.05	1.69	5.66	5.28	1.44	5.55	5.35	1.43
Financials	4.88	5.26	1.80	4.70	5.24	1.59	4.36	5.20	1.63	4.15	5.36	1.55
Technology	12.06	7.68	6.20	10.60	8.62	3.79	10.05	8.32	2.98	9.15	7.76	2.36

6. Conclusion

The analysis presented above implies regression analysis on a small sample of Australian-listed energy stocks will continue to generate risk estimates that are unreliable. This conclusion is implied by analysis of risk estimates for the Capital Asset Pricing Models and the Fama & French 3-factor model.

We consider reliability of risk estimates along two dimensions. There is the dispersion of average risk estimates across samples of firms in the same industry group, and there is the dispersion of risk estimates over time.

With respect to the first measure of dispersion, we document the extent to which risk estimates are likely to vary, depending upon which set of randomly-selected firms from the same industry we happen to observe. Specifically, if we compile beta estimates for samples of nine firms using 10 years of recent returns, the standard error of the beta estimate ranges from 0.15 to 0.22 across different industries (Table 2, Panel F). This means there is every chance that a different sample of nine firms in the same industry will produce a very different beta estimate. If we apply the same analysis to the risk premium implied by the Fama & French model, the standard error of risk premiums across six industries ranges from 1.37% to 2.48% (Table 6, Panel F).[10]

We then observe that this dispersion decreases considerably if we increase sample size. If analysis is performed on 18 firms the dispersion of average risk estimates is reduced by about one-third, and if we extend the analysis to 27 firms the dispersion of average risk estimates is reduced to about one-half. If we rely upon 27 firms, the standard error of beta estimates ranges from 0.09 to 0.12 across industries, and the standard error of Fama & French risk premiums ranges from 0.29% to 1.26%.

With respect to the second measure of dispersion, we document the extent to which risk estimates are likely to vary from one time period to another. Even if the true risk of firms varies from one period to the next, and even if this change in risk is captured by regression analysis of stock returns, substantial variation in risk estimates over time would make this technique unreliable for forecasting risk. If we consider beta estimates computed over ten year periods, the average difference in beta estimates between two 10-year periods is around 0.24 to 0.62 (Table 4, Panel B) and the average difference in Fama & French risk premium estimates is around 2.90 to 4.58% (Table 8, Panel B). The differences are higher still if only five years of returns are available for analysis.

This time-series variation in estimates is reduced if we expand sample size, but is still material. If we extend sample size to 27 firms, from the first to second ten-year period the average beta estimate changes by 0.20 to 0.58, and the average Fama & French risk premium changes by 1.53% to 3.81%.

The implication of this analysis is that a substantially more than nine firms need to be considered if we are to have any confidence that regression analysis of past stock returns will provide useful information on systematic risk. Furthermore, the results from a broader number of firms should be evaluated in conjunction with other relevant evidence both on risk estimates and the cost of equity directly. This is because regression-based estimates of risk vary across different samples of firms in the same industry, and across time for the same samples. This variation can alter the cost of capital estimate by several percentage points which has implications for investor confidence and the regulated firms' opportunity to recover at least efficient costs.

---

[10] The risk premium estimates are made after assuming values for *MRP* of 6.00%, *SMB* of 0.61% and *HML* of 2.34%. As mentioned previously, these figures are used for illustrative purposes and do not necessarily represent the best estimates of risk exposure under the Fama & French model. The *MRP* assumption of 6.00% corresponds to the current assumption of the AER and the *SMB* and *HML* assumptions are the long-term average values of the indices used to construct the monthly *SMB* and *HML* factor returns.

7. References

Australian Energy Regulator, 2009. "Electricity transmission and distribution network service providers: Review of the weighted average cost of capital (WACC) parameters – Final Decision," May.

Fama E., and K. French, 1993. "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics,* 33, 3–56.

8. Terms of reference and qualifications

This report was prepared by Professor Stephen Gray, Dr Jason Hall, Professor Robert Brooks and Dr Neil Diamond. Professor Gray, Dr Hall, Professor Brooks and Dr Diamond have made all they enquiries that they believe are desirable and appropriate and that no matters of significance that they regard as relevant have, to their knowledge, been withheld.

Professor Gray, Dr Hall, Professor Brooks and Dr Diamond have been provided with a copy of the Federal Court of Australia's "Guidelines for Expert Witnesses in Proceeding in the Federal Court of Australia." The Report has been prepared in accordance with those Guidelines, which appear in the terms of reference.

Stephen Gray and Jason Hall, SFG Consulting
Neil Diamond and Robert Brooks, Monash University