Estimating the Utilisation of Franking Credits through the Dividend Drop-Off Method

Basic statistical diagnostics and alternative models

Stefan Mero, Rohan Sadler and Richard Begley

Secretariat Working Paper

9 February 2017

Economic Regulation Authority
WESTERN AUSTRALIA

Economic Regulation Authority

4th Floor Albert Facey House
469 Wellington Street, Perth

Mail to:
Perth BC, PO Box 8469
PERTH WA 6849

T: 08 6557 7900

F: 08 6557 7999

E: records@erawa.com.au

W: www.erawa.com.au

National Relay Service TTY: 13 36 77
(to assist people with hearing and voice impairment)

We can deliver this report in an alternative format for
those with a vision impairment.

Contents

Tables

Figures

Executive Summary

Gamma is a parameter which accounts for that part of the return on equity that is realised through imputation credits an investor receives on their tax return. As a general rule, investors who are able to utilise franking credits will accept a lower required rate of return, before personal tax, on an investment that has franking credits, compared with an investment that has similar risk and no franking credits, all other things being equal.

The Economic Regulation Authority (**ERA**) in its decisions has considered that the benefit arising from imputation credits, gamma, can be interpreted as the proportion of franking credits that are distributed, multiplied by the proportion of these, theta, that are utilised by the representative investor.

The ERA in recent decisions has considered three different approaches to estimating the utilisation rate, theta:

- the equity share approach;
- the taxation statistics approach; and
- the dividend drop-off method.

This Secretariat Working Paper is concerned with the statistical aspects of the dividend drop-off method. It explores some recent criticisms of the ERA's approach to dividend drop-off estimation, and is intended to advance the debate in this area.

Specifically, the validity of the dividend drop-off (**DDO**) approach has been an important focus for recent debates about valuing gamma in a regulatory setting. For example, in its 2013 Rate of Return Guidelines, the ERA relied on the 2013 Vo, D., Gellard, B., Mero, S. conference paper on 'Estimating the Market Value of Franking Credits', when establishing its range for theta to apply in its gas decisions. This paper concluded that the 'state-of-the-art' theta estimates of the time – which were based on the models originally used by SFG Consulting in a series of estimations undertaken in 2012 – were highly unstable.

Subsequently, in March 2014, SFG Consulting was retained and instructed by Aurizon Network to provide its view on issues relating to the estimation of gamma, including those relating to the stability of dividend drop-off estimates of theta, which were raised in Vo et al.'s study. This SFG Consulting report contended that the analysis of Vo et al was 'non-standard', and hence flawed, because:

- the stability analysis did not apply a correction for broader market movements to the data; and
- it implemented a stability analysis where data observations believed to be influential were removed one-at-a-time, instead of in pairs.

Further, in December 2014, DBP submitted to the ERA a report prepared by SFG Consulting in response to an invitation for submissions on the ATCO Gas Draft Decision. This report raised the same technical issues related to the Vo et al. stability analysis as had been raised in SFG's report prepared for Aurizon in March 2014. SFG also emphasised that it had submitted its results to an expanded set of stability tests.

Most recently, in March 2016 Frontier Economics submitted a further report to the ERA on issues in relation to the regulatory estimate of gamma on behalf of Dampier Bunbury Pipeline. This report presented similar material to that set out in SFG Consulting's previous reports. Specifically, it raised the same technical issues related to the Vo et al. stability analysis and again emphasised that SFG's theta estimates are reliable and stable in light of the tests undertaken in a report for APA Group.

This study examines the criticisms of Vo et al. made by SFG Consulting/Frontier, finding that:

- generally, all theta estimates based on the market corrected data tend to fluctuate around a lower value, with a smaller range, than estimates based on uncorrected data;
- however, SFG's assertion – that a failure to apply the market correction to the data results in relatively unstable theta estimates – is only minimally supported by the analysis;
    - the scale of fluctuation in the results, with the market corrected data, tends to remain large relative to the value of the theta estimate;
- accordingly, the results with the market corrected data are not 'stable', as they change considerably, depending on the model and regression technique chosen;
- further, removing influential observations in pairs, as is done by SFG, instead of one-at-a-time, as is done by Vo et al, 'induces' smoothness, which falsely increases the appearance of stability in the theta estimates.

More importantly, while SFG/Frontier have paid considerable attention to these stability analyses, they have done so without considering more commonly used econometric diagnostics. This study applies these econometric diagnostics to SFG/Frontier's models. The results of the diagnostics show that SFG/Frontier's models are a very poor fit on the data. Applying better fitting models can produce much higher estimates of theta than those proposed by SFG/Frontier. Despite finding superior fitting models, all dividend drop-off models still fit the data very poorly, rendering theta estimates from the dividend drop-off method highly uncertain.

Further Information

Further information regarding this report can be obtained from:

Stefan Mero
Regulation and Inquiries Division
Economic Regulation Authority
Ph (08) 6557 7900

1 Introduction

1. Gamma is a parameter in the Weighted Average Cost of Capital that accounts for that part of the return on equity that is realised through imputation credits an investor receives on their tax return. As a general rule, investors who are able to utilise franking credits will accept a lower required rate of return, before personal tax, on an investment that has franking credits, compared with an investment that has no franking credits (all other things being equal).

2. The Economic Regulation Authority (**ERA**) estimates gamma (γ) as the product of the distribution rate F and the estimate of the utilisation rate θ (theta):[1]

$$\gamma = F \times \theta \tag{1}$$

3. The benefit arising from imputation credits is interpreted as the proportion of franking credits distributed (F) multiplied by the proportion of these that are utilised by the representative investor (θ).[2] The ERA's interpretation is consistent with that of the Australian Energy Regulator, which describes the utilisation rate as 'the utilisation value to investors in the market per dollar of imputation credits distributed'.[3]

4. The ERA in recent decisions has considered three different approaches to estimating the utilisation rate, theta:[4]

   - the equity share approach;
   - the taxation statistics approach; and
   - the dividend drop-off method.

5. This paper is concerned with the statistical aspects of the dividend drop-off method for estimating the utilisation rate.[5] Issues relating to whether dividend drop-off studies correctly estimate theta as defined are not examined in this statistical review. The ERA has observed that dividend drop-off studies may not correctly estimate the utilisation rate. This means that theta (θ), which is estimated in the dividend drop-off framework set out in this study, *is not necessarily the same theta* that is expressed in equation (1).

6. Dividend drop-off studies observe the change in the prices of stocks when they go from trading 'cum-dividend', where the current holder is entitled to the dividend, to

---

[1] This follows the analysis by Monkhouse in relation to the impact of imputation credits on the effective tax rate of companies. See equation 2.5 in P. Monkhouse, The valuation of projects under the dividend imputation tax system, *Accounting and Finance*, 36, 1996, p. 192; Goldfields Gas Pipeline, *Access Arrangement Revision Proposal: Supporting Information*, 15 August 2014, Appendix 1.

[2] Economic Regulation Authority, *Final Decision on Proposed Revisions to the Access Arrangement for the Dampier to Bunbury Natural Gas Pipeline 2016 – 2020*, Appendix 5 – Gamma, 30 June 2016, p. 10.

[3] Australian Energy Regulator, *AusNet Services distribution determination final decision 2016–20*, Attachment 4, p. 9.

[4] For a detailed consideration of each approach, see Economic Regulation Authority, *Final Decision on Proposed Revisions to the Access Arrangement for the Dampier to Bunbury Natural Gas Pipeline 2016 – 2020*, Appendix 5 – Gamma, 30 June 2016.

[5] See Economic Regulation Authority, *Final Decision on Proposed Revisions to the Access Arrangement for the Dampier to Bunbury Natural Gas Pipeline 2016 – 2020*, Appendix 5 – Gamma, 30 June 2016, pp. 11-12).

trading 'ex-dividend', where any new holder of the stock is no longer entitled to the dividend yet to be paid. The idea behind these studies is that the observed change in price should reflect the market's value attached to the gross dividend being paid.

7. However, the 'gross' dividend consists of both cash (or 'net') dividends and company tax paid on the gross dividend. The franked proportion of gross dividends return the tax paid to investors in the form of imputation credits. Econometric regression techniques are employed to quantify the change in price across a sample of stocks and estimate the quantity of the change that can be attributed to net dividends and imputation credits. Theta (θ) is the proportion of price change attributable to gross dividend payment that is in turn, attributed to imputation credits.

8. In its Rate of Return Guidelines the ERA relied on dividend drop-off studies undertaken by both Vo et al and SFG Consulting to establish a permissible range for theta of 0.35-0.55.[6] Vo et al.'s dividend drop-off study raised issues relating to the stability of estimates produced by the technique employed by SFG. Vo et al.'s study tested stability using a method that sequentially removes an increasing number of influential data points, re-estimating theta each time a data point is removed. The results indicated that dividend drop-off estimates of theta are sensitive to inclusion or exclusion of a small number of highly influential observations.

9. In March 2014, SFG Consulting was retained and instructed by Aurizon Network to provide its view on issues relating to the estimation of gamma, including those relating to the stability of dividend drop-off estimates raised in Vo et al.'s study.[7] This report proposed that the analysis of Vo et al was non-standard because:

- the stability analysis did not apply a correction for broader market movements to the data; and
- it implemented a stability analysis where data observations believed to be influential were removed one-at-a-time, instead of in pairs.[8]

10. The market correction issue relates to the preparation of the price data used in the dividend drop-off analysis. Some of the change between cum-and ex-dividend price will stem from movement in the stock market as a whole. To factor the market movement out of the price change, market returns are discounted out of ex-dividend prices. SFG Consulting's report to Aurizon, to some extent, attributes Vo et al.'s finding of instability in estimates to the omission of applying a market correction to the data.[9]

11. The issue relating to pairwise removal of observations concerns the application of the stability test in Vo et al. As mentioned above, the Vo et al study tested stability by removing an increasing number influential data points, and re-estimating theta each time a data point is removed. This method begins by detecting the most influential observation through a DFBETAs statistic, remove this observation from the sample,

---

[6] Economic Regulation Authority, *Explanatory Statement for the Rate of Return Guidelines*, 16 December 2013, pp.209-223.
Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013.
SFG Consulting, *Dividend drop-off estimate of theta*, Final Report, 21 March 2011.

[7] SFG Consulting, *Estimating Gamma: Report for Aurizon Network*, 6 March 2014, p.1.

[8] Ibid, pp.10-11.

[9] Ibid, p.11.

and then re-estimate theta. This process is repeated one observation at a time until 30 observations were removed.

12. SFG have implemented tests that remove outliers and then re-estimate theta in a number of different ways. One of their methods for testing stability involves removing observations in pairs instead of one-at-a-time. The pairs consist of observations where the first observation has the most influential upward effect on the estimate of theta, and the second observation has the most influential downward effect on the estimate. SFG Consulting's report to Aurizon concludes that the finding of instability in estimates is likely to be manifest in the Vo et al. 'non-standard' approach applied.[10]

13. In December 2014, DBP submitted a report prepared by SFG Consulting to the ERA in response to an invitation for submissions on the ATCO Gas Draft Decision. This report raised the same market correction issue discussed in SFG's report prepared for Aurizon in March 2014. In relation to the stability analysis, SFG emphasised that it had subjected its results to an expanded set of stability tests. It appears that SFG was making reference to its report prepared in May 2014 on behalf of APA, and submitted as part of the proposed revisions to the Goldfields Gas Pipeline access arrangement. In this report, SFG submitted that it applied the 'one-at-a-time' influential observation approach that Vo et al. used in their study and in addition conducted a randomised bootstrapping analysis. This analysis randomly eliminates 5 per cent of the data and re-estimates theta. The procedure is repeated 1000 times to produce a distribution for theta estimates.[11] SFG concludes that this analysis corroborates the stability and insensitivity of its theta estimates to the removal of outliers.

14. In March 2016 Frontier Economics submitted a report on issues in relation to the regulatory estimate of gamma behalf of DBP. This report submitted the material presented in SFG Consulting's previous reports. Specifically, it raised the market correction issue and again emphasised that SFG's theta estimates are reliable and stable in light of the tests undertaken in its report for APA in May 2014.[12]

15. This study seeks to address the issues originally raised in SFG's report prepared for Aurizon in March 2014.[13] The impact of the market correction on Vo et al.'s sensitivity analysis is assessed by comparing the original results without the market correction to results based on the original data that include the market correction. It demonstrates that there are other statistical issues affecting certainty around theta estimates that are more important than the market correction.

16. The effect of perceived 'outliers' on estimates of theta have been a main point of contention. This study assesses the value of sensitivity analyses which repeatedly remove what are believed to be outliers and re-estimate theta. More meaningful tests are then applied to the dividend drop-off study methods employed by both Vo et al and SFG to determine whether the methods violate standard statistical assumptions. This study then moves on to demonstrate that, in the face of non-normally distributed data, model specification combined with choice of estimation method is the primary

---

[10] Ibid, p.11.

[11] SFG Consulting, *An appropriate regulatory estimate of gamma: Report for Jemena Gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks*, 21 May 2014, pp. 95-101.

[12] Frontier Economics, *Issues in relation to the regulatory estimate of gamma: A report prepared for DBP*, March 2016.

[13] SFG Consulting, *Estimating Gamma: Report for Aurizon Network*, 6 March 2014, p.1.

driver of differences and thus uncertainty in dividend drop-off analysis results. The issue of assessing the fit of a specified model is addressed and an estimate of theta based on a better fitting model is produced.

2 The Dividend Drop-off Study Framework

17. The basic model employed in dividend drop-off studies is shown in equation (2).

$$\Delta P_i = \delta Net\ Dividend_i + \theta Franking\ Credit_i + \varepsilon_i \qquad (2)$$

Where:

ΔP is the change in the price across the cum-and ex-dividend date of a given stock i;

δ is the change in ΔP per dollar change in net dividend;

$Net\ Dividend_i$ is the size of the net dividend credited to the holder of stock i at a given on the cum-dividend date;

θ (theta) is the change in ΔP per dollar change in franking credit;

$Franking\ Credit_i$ is the size of the franking credit that is credited to the holder of stock i on the cum-dividend date; and

$\varepsilon_{i,t}$. is the error between the regression prediction and observation on stock i.

18. This equation expresses the change in stock prices between the cum-and ex-dividend date as reflecting the net or 'cash' dividend, franking credit measured on observed data with some random error (which is zero on average). This model is typically implemented in a statistical regression framework. Ex-and cum-dividend date price data is observed on a number of stocks across a number of dividend 'payment' events for each stock through time. This is then regressed on the corresponding observations of net dividend and franking credit data in order to estimate the parameters δ and θ which are inferred to be the 'market' value of the net dividend end and franking credit, respectively.[14]

19. The standard regression method is ordinary least squares (**OLS**). OLS relies on the assumption, among others, of a constant variance in the error terms $\varepsilon_{i,t}$ (no heteroskedasticity). When this assumption is violated the estimated standard errors from the model used to test the statistical significance of estimated parameters tend to be too small, resulting in a failure to correctly conclude parameters are not statistically different from zero (type I error). The price data used in the dividend drop-

[14] As noted above, issues relating to whether dividend drop-off studies correctly estimate theta as defined are not examined in this statistical review. The ERA has observed that dividend drop-off studies may not correctly estimate the utilisation rate (Economic Regulation Authority, *Final Decision on Proposed Revisions to the Access Arrangement for the Dampier to Bunbury Natural Gas Pipeline 2016 – 2020*, Appendix 5 – Gamma, 30 June 2016, pp. 11-12).

off studies tend to exhibit non-constant variance (heteroskedasticity). This means the estimated parameters δ and θ may produce estimates that appear statistically different from zero when in fact, this is not the case. Various reasons for the observed heteroskedasticity have been posited, however it appears no consensus has been reached as to the exact causes.[15]

20. SFG's approach attempts to address the heteroskedasticity issue by specifying four additional forms of the model shown in equation (2).[16] This approach scales the observations of variables used in the model by dividing them through variables believed to be related to potential 'patterns' observed in the (non-constant) variance, thereby removing those patterns to produce constant variance.

21. The four models are shown below.

$$\frac{\Delta P_i}{P_{Cum,i}} = \delta \frac{Net\ Dividend_i}{P_{Cum,i}} + \theta \frac{Franking\ Credit_i}{P_{Cum,i}} + \varepsilon_i \qquad (3)$$

22. The model shown in equation (3) assumes variance is related to cum-dividend prices. Dividing each term by cum-dividend price expresses all of the variables in in the regression terms of 'yield' on the cum-dividend price. This is referred to as **model 1** in this study.

$$\frac{\Delta P_i}{Net\ Dividend_i} = \delta + \theta \frac{Franking\ Credit_i}{Net\ Dividend_i} + \varepsilon_i \qquad (4)$$

23. The model shown in equation (4) assumes variance is related to the net dividend. Note the net divided disappears from the right side of the equation as a result of being divided by itself. This is referred to as **model 2** in this study.

$$\frac{\Delta P_i}{Net\ Dividend_i \sigma_i} = \frac{\delta}{\sigma_i} + \theta \frac{Franking\ Credit_i}{Net\ Dividend_i \sigma_i} + \varepsilon_i \qquad (5)$$

24. The model shown in equation (5) assumes that variance is related to the standard deviation of stock i's returns in excess of the market return (σ_i).[17] This is referred to as **model 3** in this study.

---

[15] The size of variance in the error terms may be related to the net dividend and franking credit, but not detected as having a relationship with the price drop-off because they are related to *absolute size* of the drop-off as opposed to the *sign* of the drop-off. Another hypothesis is that larger companies tend to trade more frequently and so have lower variance in their error terms that can stem from sporadic pricing in less traded smaller stocks. See Beggs and Skeels, 'Market Arbitrage of Cash Dividends and Franking Credits', *The Economic Record*, vol.82, no.258, 2006, p.243.

[16] SFG Consulting, *Dividend drop-off estimate of theta*, Final Report, 21 March 2011, pp. 19-21.

[17] $\sigma_{i,t} := \sqrt{\frac{1}{N}\sum_{j=1}^{N}(er_{i,t-5-j} - \overline{er}_{i,t})}$ is the estimated standard deviation of excess returns of stock i over N trading days. $\overline{er}_{i,t} := \frac{1}{N}\sum_{j=1}^{N}\overline{er}_{i,t}$ is the estimated mean of excess return of stock i over N trading days. $er_{i,t} := r_{i,t} - r_{m,t}$ is the excess return of stock i over the market at time t. $r_{i,t}$ is the return of stock i at time t. $r_{m,t}$ is the return of the All Ordinaries Index at time t. More details are given in Appendix 1.

$$\frac{\Delta P_i}{P_{Cum,i}\sigma_i} = \delta \frac{Net\ Dividend_i}{P_{Cum,i}\sigma_i} + \theta \frac{Franking\ Credit_i}{P_{Cum,i}\sigma_i} + \varepsilon_i \qquad (6)$$

25. The model shown in equation (6) assumes that variance is related to the cum-dividend date price multiplied by the standard deviation in excess returns. This is referred to as **model 4** in this study.

26. The ERA in its 2013 Rate of Return Guidelines relied upon Vo et al. in estimating parameters for all four of these specifications.

27. A number of different linear regression methods can be, and have been, used to estimate the parameters δ and θ in these models. These methods include OLS, MM robust regression (**MM**) and least absolute deviation (**LAD**) regression. The robust regression methods have been applied by SFG and Vo et al. in an attempt to mitigate the influence of what are perceived to be 'outliers' in the regression.

3 Addressing issues raised by SFG and Frontier Consulting

3.1 Market corrected data and sensitivity analysis

28. As outlined in the introduction, SFG Consulting have raised concerns that the stability analysis of Vo et al., relied upon by the ERA in its 2013 Rate of Return Guidelines, did not apply a correction for broader market price movements to the data.[18] Vo et al. applied this stability analysis to show that removing a very small number of outliers can vary the estimate of theta from 0.3 to 0.55.[19] This was still the case when robust regression techniques were applied.

29. SFG and Frontier Economics partly attributes Vo et al.'s finding of instability in estimates of theta to the omission of applying this market correction to the data.

30. The market correction of data is applied as shown in equation (7).

$$P_{Ex,i}^{Adjust} = \frac{P_{Ex,i}}{1+r_{m,i}} \qquad (7)$$

Where:

$P_{Ex,i}$ is the price of stock i on the ex-dividend day at time t; and

$r_{m,i}$ is the daily return on the All Ordinaries index on the ex-dividend date.

31. To test SFG's assertion, that a failure to apply this market correction results in instability in theta estimates, the same stability tests on the same data set in the Vo et al. analysis are undertaken below, but with the market correction applied. These results are compared to the original results below.

---

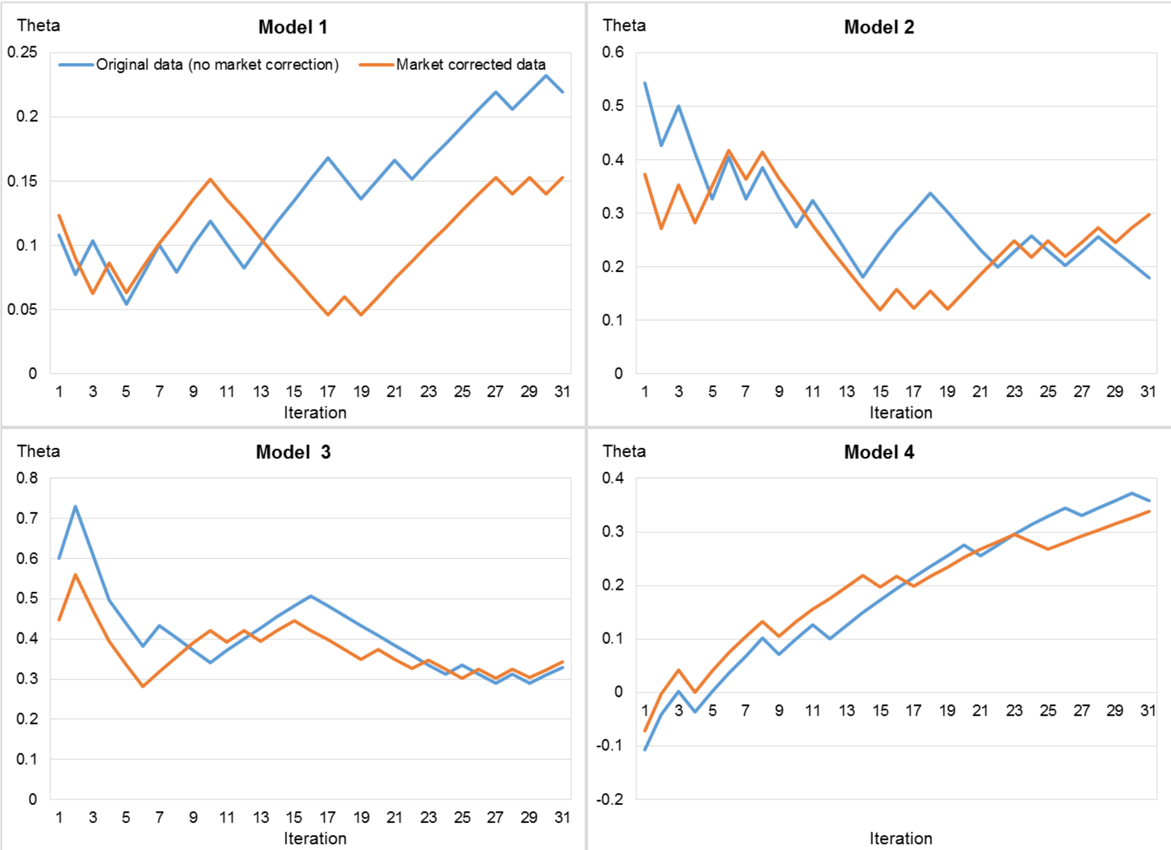
[18] Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013.

[19] Ibid, p.31

32. Figure 1 compares the results of sensitivity analysis applied in Vo et al. based on the OLS regression estimates. The sensitivity analysis removes the most influential data points and re-estimates theta 30 times to observe the consequent effect on theta. Data points are classified as influential according to the 'DFBETAs' criterion (see Appendix 2 for technical details). This is a standardised measure of the amount by which a regression coefficient, such as theta, changes if a particular observation is removed. Each iteration on the x axis is a re-estimate of theta. Each of the four models for estimating theta outlined in chapter 2 are shown in their respective panels.

33. These results include those based on market corrected data, which tend to fluctuate within a reduced range, as compared to those based on uncorrected data. The estimates based on market corrected data also tend to fluctuate around a lower estimate of theta than the estimates based on the original data.

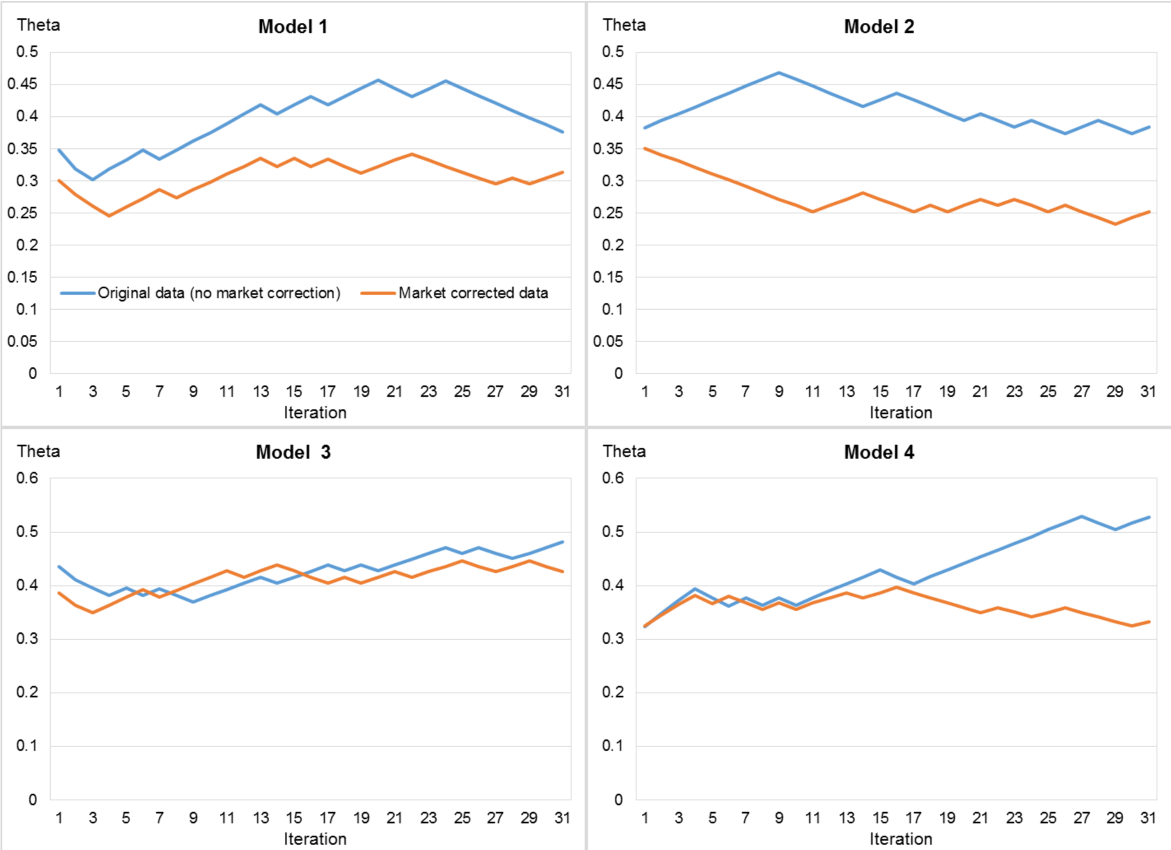
Figure 1 Sensitivity analysis of OLS theta estimates – with and without market correction



Source: ERA Analysis

34. Figure 2 compares the results of the MM regression estimates. In this case, with the exception of model 2, the results based on market corrected data tend to fluctuate within a reduced range. The estimates based on market corrected data, again, also tend to fluctuate around a lower estimate of theta than the estimates based on the original data.

Figure 2 Sensitivity analysis of MM theta estimates – with and without market correction



Source: ERA Analysis

35. Figure 3 compares the results of the LAD regression estimates based on data with no market correction to results based on data that incorporates the market correction. Vo et al. did not report LAD stability test results for models 3 and 4.

Figure 3 Sensitivity analysis of LAD theta estimates – with and without market correction

Source: ERA Analysis

36. The conclusion drawn from Figure 1 and Figure 2 can be also be drawn from Figure 3. The results based on market corrected data tend to fluctuate within a reduced range and around a lower estimate of theta.

37. Application of the market correction to the data seeks to control for market wide effects in the drop-off. However, SFG's assertion that a failure to apply the market correction to the data results in instability in theta estimates is only minimally supported by the analysis above.

38. With the exception of the model 2 MM estimate, all theta estimates based on the market corrected data tend to fluctuate around a smaller range than estimates based on uncorrected data. However, the scale of fluctuation tends to remain large relative to the value of the theta estimate.

39. Moreover, it is evident from the figures above that the results are not 'stable' from the perspective that they change considerably depending on the model and regression technique chosen. The most 'stable' estimate in the analysis above is the LAD estimate based on model 2 (Figure 3), however, changing from model 2 to model 1 results in greater fluctuation in the estimate around a lower level of theta. The fluctuation increases with MM and OLS estimates. This raises two questions. Firstly, what does the sensitivity analysis tell us? Secondly, what model and regression technique is the most appropriate?

3.2 What do the sensitivity analyses tell us?

40. The sensitivity analyses applied by SFG and subsequently Vo et al. involve removing outliers and re-estimating theta to observe the consequent effect on theta. These tests have been implemented in a number of ways including:

- comparing estimates of theta before and after removing the most influential 1 per cent of observations; [20]
- pairwise removal; [21]
- one-at-a-time removal; [22] and
- bootstrapping (simulating distributions) removing 5 per cent of data.[23]

41. Comparing estimates of theta before and after removing the most influential 1 per cent of observations uses 'Cook's distance' to measure and rank the influence of observations in the data set on the final parameter estimates. Cook's distance measures the effect of removing a given observation on estimates of the regression parameters. The most influential one per cent of observations in terms of Cook's distance are removed from the data set (all at once) and theta is re-estimated.

---

[20] SFG Consulting, *The impact of franking credits on the cost of capital of Australian companies: Report prepared for Envestra, Multinet and SP Ausnet*, 25 October 2007.

[21] SFG Consulting, *Dividend drop-off estimate of theta*, Final Report, 21 March 2011, p.28.

[22] SFG Consulting, *An appropriate regulatory estimate of gamma: Report for Jemena Gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks*, 21 May 2014, p. 95.
Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, p.26.

[23] SFG Consulting, *An appropriate regulatory estimate of gamma: Report for Jemena Gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks*, 21 May 2014, p. 98.

42. Pairwise removal of observations first determines which single observation, if removed, would result in the greatest increase in the estimate of theta (see Appendix 2 on SFG and Frontier's stability analysis for further discussion). Next it determines which single observation, if removed, would result in the greatest decrease in the estimate of theta. Both observations are then removed and theta is re-estimated. This process is repeated by removing another pair of observations re-estimating theta. This continues until 25 pairs of observations have been removed in total. The parameter estimates are then plotted sequentially on a graph in the same way as shown in section 3.1.

43. As shown and discussed in section 3.1, one-at-a-time removal of observations uses the DFBETAs criterion to identify influential data points instead of other measures such as Cook's distance. DFBETAs is a standardised measure of the amount by which a regression coefficient changes if a particular observation is removed.[24] The parameter estimates are then plotted sequentially on a graph in the same way as shown in section 3.1.

44. Bootstrapping simulates a distribution for the theta parameter. A random five per cent of the data sample is eliminated, with theta re-estimated from the remaining 95 per cent of the observations. This process is repeated 999 times to create 1000 estimates of theta (by including the full sample estimate). The resultant bootstrapped estimates are plotted as a histogram to show the distribution of the estimates.

45. The tests based on one-at-a-time and pairwise removal of observations described above are typically employed to address concerns that a time series model's parameters may be different during a forecast period to what they are during the sample period. Use of such tests in the current context of estimating theta, which is based on observations on a cross-section of stocks, is unusual.[25] In this context the movement in the sequential plot of recursive estimates reflects removal of data that least comply with the assumed fit.

46. For OLS regression, greater movement in the stability plots will tend to indicate a poorer fitting model. This is because the OLS regression method does not down-weight extreme data points in the model fitting process and takes them into account to provide the best fit. For this reason, the removal of extreme data points will have a stronger effect on the OLS fit and theta estimate than for MM and LAD regression, which are designed to mitigate the effect of extreme data points. When the model is a poor fit to the data, removing several of the most extreme data points will tend to alter this fit and theta estimate, more so than the removal of less extreme data points. Each subsequent point removed will tend to have less of an impact on the fit (and thus theta parameter), giving the impression of stability. If a large proportion of extreme observations in the sample have been removed by a filtering process prior to this type of stability test then the plots are likely to converge and 'stabilise' more quickly.

47. The plot of robust regression estimates (MM and LAD) derived from recursively removed data will provide an impression of greater stability than OLS. This is because these regression methods down-weight, and hence mitigate, the effect of extreme data points with each sequential theta estimate. The recursive plots based on MM and LAD regression are therefore predisposed to show a lack of movement by already

---

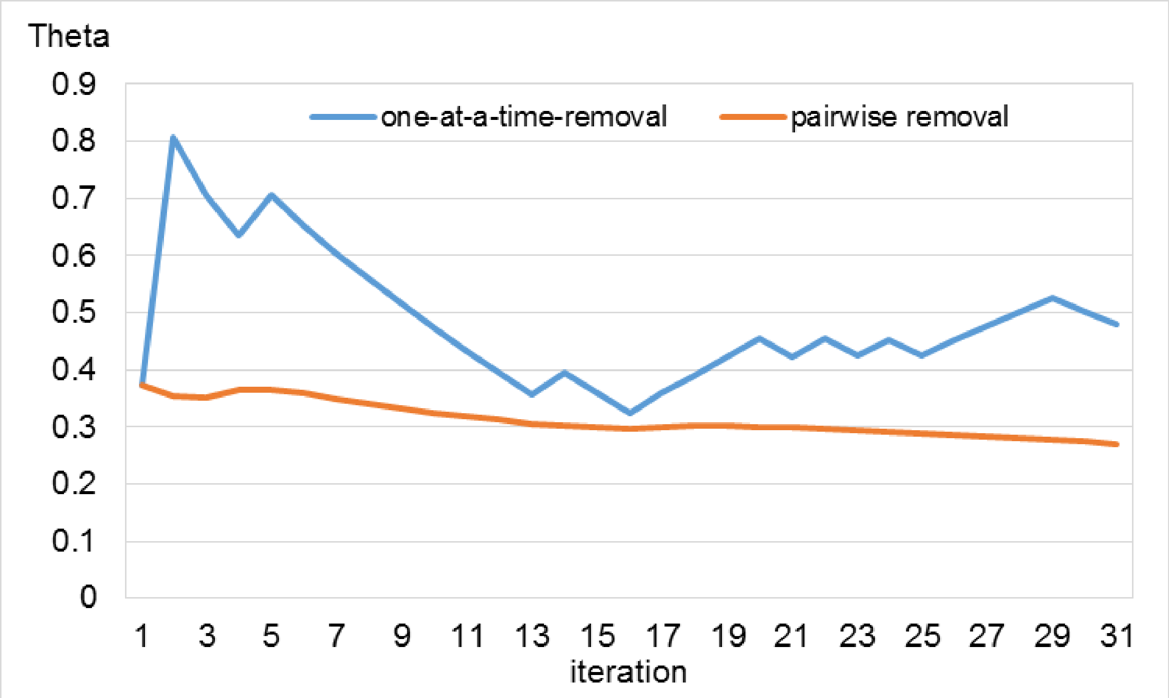
[24] Technical details are given in Appendix 2 in the section on SFG and Frontier's stability analysis.

[25] While the observations used in the theta regressions span a period of time, the models themselves do not incorporate changes in time reflected in the data.

being relatively insensitive to extreme data points. This relative insensitivity of robust regression estimates is evident in the comparison of Figure 2 and Figure 3 to Figure 1. The estimates based on OLS in Figure 1 fluctuate substantially, while the MM and LAD estimates in Figure 2 and Figure 3 fluctuate less so.

48. Removing pairs of observations with offsetting impacts on theta tends to 'smooth' the stability plots. The effect of using pairwise removal versus one-at-a-time removal is shown in Figure 4. More technical details regarding the smoothing effect of pairwise removal can be found in Appendix 2 in the section 'Critique of the DFBETAs Stability Analysis'.

Figure 4 Pairwise vs one-at-a-time removal in stability tests on model 3 OLS estimates



Source: ERA Analysis

49. This apparent stability is similarly observed when comparing estimates of theta before and after removing the most influential 1 per cent of observations using Cook's distance. Removing extreme observations is likely to assist in stabilising estimates.[26]

50. In short, these methods induce 'stability' through removing or mitigating the effect of non-compliant observations using robust regression or removing pairs of observations with offsetting effects. Instead of specifying a model that complies with the data the focus of these analyses has been to mitigate the effect of non-compliant data to create similar looking sets of theta estimates. Because of this, the similar or 'stable' looking sets of theta estimates provide little information about certainty around parameter estimates or robustness of the estimates to possible violations of other basic modelling assumptions.

51. Bootstrapping by randomly removing 5 per cent of observations and re-estimating theta differs from the other sensitivity analyses in that it does not target influential observations. It demonstrates the sensitivity of theta to removing random

[26] This is evident in SFG's analysis of sub-periods across table 3 and 4 in SFG Consulting, *The impact of franking credits on the cost of capital of Australian companies*, 12 November 2008, pp. 2-3.

observations regardless of whether they are influential or not. In effect, removing 5 per cent of the observations slightly 'thins' the sample. The results based on the thinned sample will be expected to produce a similar estimate of theta. This is because the full sample and any large subsample of the observations are not likely to have substantially different distributions. Due to the central limit theorem, the bootstrapped distribution of either the full sample, or a thinned sample, will be approximately normal. The only difference will be that the thinned sample will have a slightly larger standard error associated with theta (see discussion in Appendix 2, 'A Comment on the SFG Bootstrap Sensitivity Analysis'). This type of bootstrapping analysis therefore says little, if anything, about the stability of estimates because it is predisposed to producing a largely smoothed and normally distributed histogram of the sampling distribution of theta.

52. All of the sensitivity analyses applied (including bootstrapping) are focussed on the behaviour of estimated parameters. The tests ignore the behaviour of discrepancies between the fitted models' predictions (residuals) and the actual data. Rather than focus on the stability of parameter estimates, conventional regression diagnostics assess the behaviour of these residuals to inform the appropriateness of the proposed model with respect to the data. That is, once a model is fitted and parameters are estimated the residuals are observed to determine if they exhibit any remaining patterns that have not been adequately captured by the model specification. The residuals are also observed to determine whether they follow the distribution assumed when specifying the model and regression fitting procedure. The emphasis of diagnostics should be focused on the model fitting process, rather than reiterating parameter estimation over different truncated samples.

4 Are the models that produce the theta estimates meaningful?

53. As mentioned above, conventional statistical diagnostics assess the behaviour of the discrepancies between the fitted model and actual data, as measured by the residuals, before settling on a model and regression technique to produce the required parameter estimates.[27]

54. Although the classical linear regression model is intuitively appealing in the sense of the drop-off being composed of a linear combination of net dividends and franking credits, few diagnostics have been undertaken to determine whether its modification with scalars (such as those applied in models 1 through to 4) results in a nonlinear relationship. If this is the case a linear model is likely to be a poor fit and associated with relatively higher uncertainty in the estimates.

55. Hence, regression diagnostics should be designed to test for violations of the classical linear regression model (**CLRM**) assumptions. These assumptions are:

- The response variable, for example, the dividend drop-off, is a linear function of the predictor variables (net dividend and theta).

- Values observed for the predictor variables (for example net dividends and franking credits) are randomly sampled.

[27] The issue of assessing residuals was indirectly raised in Lally. M, *The Estimated Utilisation Rate for Imputation Credits*, 12 December 2012, p.21.

- No multicollinearity - the predictor variables are not an exact linear function of each other.

- The residuals at any given value of the predictor variables have a mean of zero.

- No heteroskedasticity - the residuals exhibit constant variance as a function of the predictor variables.

- No correlated errors - the residuals do not exhibit a pattern across time or within different groupings of the data.

- No correlation between residuals and predictor variables.

- The regression model is correctly specified – the relationship between the predictor variables and the response variable is correctly stated by the model.

- The number of observations must be greater than the number of parameters to be estimated.

56. Testing for violations of these assumptions assists in producing meaningful parameter estimates. Such techniques involve visual analysis of the residuals and testing them via a suite of statistical procedures. Such diagnostics have received little attention in Australian theta studies so far despite being standard practice in applied econometrics.

57. To test the assumption of a linear regression being the correct specification Ramsey's RESET test is applied to each model. The test is commonly used in basic econometrics courses. An F-statistic larger than the critical value or p-value less than 0.05 indicates rejection of the hypothesis that the linear model is the correct specification at the 5 per cent significance level.[28]

Table 1 Ramsey RESET tests of model specification error

Model	F-statisitic	p-value	outcome
1	45.0042	<0.0001	Reject hypothesis model correctly specified
2	0.2877	0.75	Do not reject hypothesis model correctly specified
3	8.6396	0.0002	Reject hypothesis model correctly specified
4	231.4059	<0.0001	Reject hypothesis model correctly specified

Source: ERA Analysis

Note: the resettest function in R package lmtest yield identical results for MM regressions

58. The results indicate that a linear specification is only appropriate for model 2 which is scaled by net dividends. This suggests that nonlinear patterns in the raw drop-off and franking data are adjusted out by dividing the data by net dividends. The results indicate that it is inappropriate to apply classic linear regression techniques to the other models because they produce a relatively poor fit compared to nonlinear estimates. This issue is examined in further detail in the discussion of robust models in Appendix 2. Other regression techniques, however, may produce more meaningful estimates for models 1, 3 and 4.

[28] The conventional default assumption of adding a square and cubic term to the alternate specification is applied.

59. The assumption of variability and random sampling in the predictor variables may be violated by systematic removal of observations from the dividend events observed over a given period, by applying filtering techniques or removing observations based on some other criteria. The issue of removing extreme data points has been highlighted at length in section 3.2. Violation of this assumption potentially induces bias in estimates.

60. Multicollinearity is an inherent issue in dividend drop-off studies. The net dividend and franking credit are highly correlated and display a strong linear relationship. This is because in the majority of cases the franking credit equals 30 per cent (corporate tax rate) of the gross dividend while the net dividend comprises the remaining 70 per cent. In regression analysis using multiple variables, multicollinearity can obscure the true relationship between each individual predictor variable and the dependent variable while the predictive accuracy of the whole regression remains valid. In addition, the standard errors on the affected parameters tend to be inflated, which results in incorrectly concluding that parameters are not statistically different from zero (type II error). For example, the situation could arise where one *incorrectly* concludes that theta is not significantly different from zero.

61. The linear relationship between net dividends and franking credits is not perfect in the current context because it is obscured by differences in the proportion of gross dividends that are franked. That is, the franking credit in some instances is less than 30 per cent of the gross dividend due to less than 100 per cent of the gross dividends being eligible for franking credits. Consequently, the issue of multicollinearity does not completely invalidate the DDO study, but is a source of increased uncertainty in the estimates. Further details on the extent to which this is an issue are covered in the discussion of multicollinearity in Appendix 2.

62. Variance in the residuals of each model may be dependent on the level of franking credits. Such a relationship violates the assumption of no heteroskedasticity. A formal test of this is the Breusch-Pagan test. The Breusch-Pagan test applied here tests for a relationship between the explanatory variables in each model (such as net dividends and franking credits) and variance in the residuals such as that shown in Figure 5. The null hypothesis is that there is no relationship. A large Breusch-Pagan statistic relative to the critical value (not shown here), or small p-value, indicates rejection of the null hypothesis of no relationship and we conclude that heteroskedasticity is present.

Table 2 Breusch-Pagan Test for Heteroskedasticity in residuals from OLS/MM regressions

Model	Breusch-Pagan Statistic	p-value	outcome
1	3.09	0.08	Do not reject hypothesis of no heteroskedasticity
2	0.01	0.94	Do not reject hypothesis of no heteroskedasticity
3	1.06	0.30	Do not reject hypothesis of no heteroskedasticity
4	60.27	0.00	Reject hypothesis of no heteroskedasticity

Source: ERA Analysis

Note: Results are identical between OLS and MM regressions using bptest in R package lmtest.

63. In Table 2 model 4 reports a large test statistic with a p-value less than 0.05. This indicates rejection of the hypothesis of no heteroskedasticity at the 5 per cent significance level.

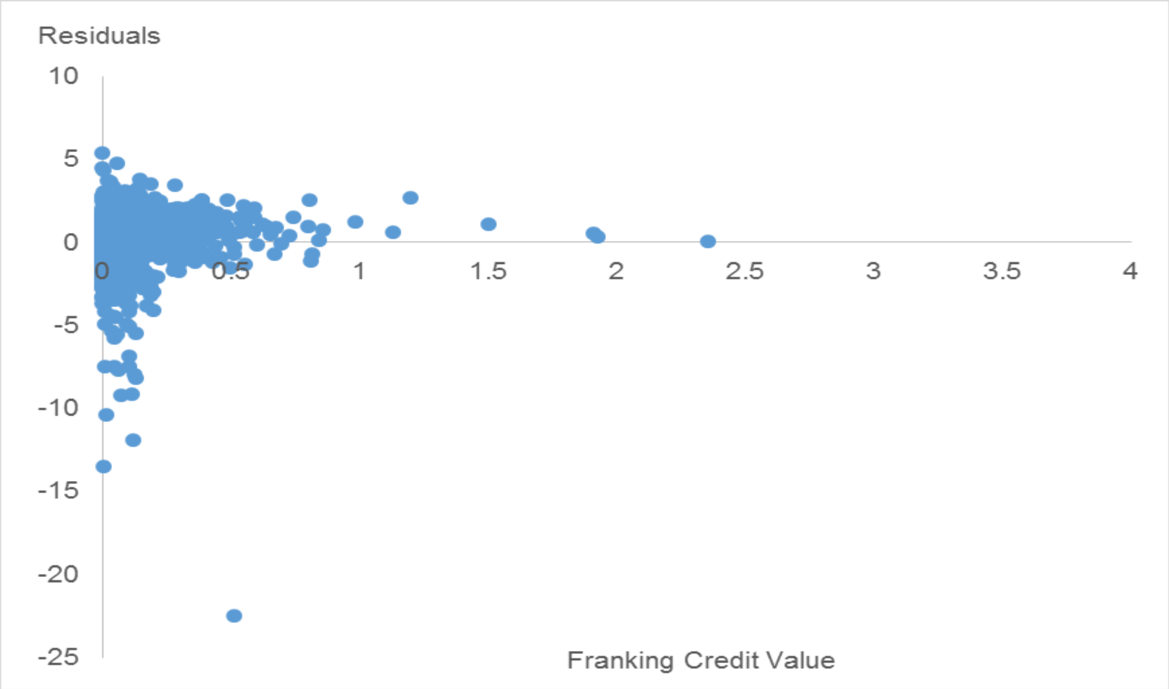
Figure 5 **Model 4 MM regression residuals plotted against corresponding franking credit values (market corrected data)**

Source: ERA Analysis

Note: The plot has truncated a single observation for a franking value credit valued between 9 and 10 in order to show the pattern in the bulk of the data. In the context of the model 4 regression franking credit values would be of a different scale due to being divided by cum price multiplied by the standard deviation in excess returns.

64. Figure 5 shows a distinguishable relationship between the residuals and franking credit value. Lower values of franking credits are associated with a higher discrepancy between the model predictions and the actual data than lower values. This indicates that a relationship between franking credits and variance in parameter estimates exists in the residuals which is not adequately captured by model 4.

Figure 6 **Model 4 OLS regression residuals plotted against corresponding franking credit values (market corrected data)**



Source: ERA Analysis

Note: The plot has truncated a single observation for a franking value credit valued between 9 and 10 in order to show the pattern in the bulk of the data.

65. Figure 6 based on the OLS regression residuals of model 4 indicates an almost identical relationship. Again, this confirms that a relationship between franking credits and variance in parameter estimates exists which is not adequately captured by model 4.

66. The consequence of this heteroskedasticity in the residuals of model 4 is that estimated standard errors from the model used to test the statistical significance of estimated parameters tend to be too small. The smaller standard errors produced by the model were one of the factors behind SFG's preference for model 4. [29] The presence of heteroskedasticity in the residuals generated by the model indicates that the small standard errors are unreliable and that model 4 is not a wholly valid specification. Further details on this issue are given in the discussion of heteroskedasticity in Appendix 2.

67. Figure 2 shows model 4 as being relatively stable. SFG also expressed a preference for this model on the basis of its estimate of theta being 'very stable'. [30] The diagnosis of residuals for heteroskedasticity shows that the stability tests are an inadequate test of model validity.

68. Aside from violations of the CLRM assumptions, an important aspect to note in the models fitted in the dividend drop-off studies is that the residuals are not normally distributed. This is a fundamental assumption behind evaluating the statistical significance of parameters in models fitted to a finite data set.

---

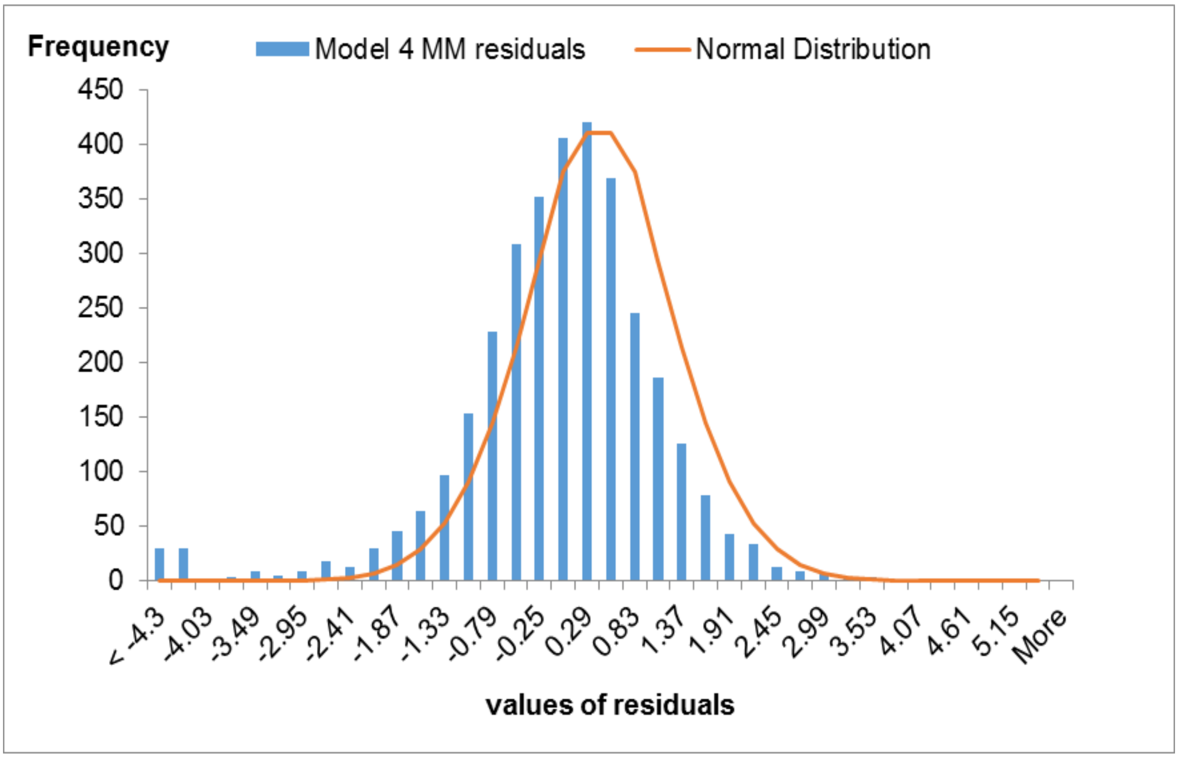
[29] SFG Consulting, *Dividend drop-off estimate of theta: Final Report*, 21 March 2011, pp.31-32.

[30] Ibid.

69. Some deviations from a normal distribution are to be expected and are acceptable, given the distribution of samples from a normally distributed population will often differ from the population idiosyncratically. Significant deviations from a normal distribution, however, can render the standard errors and thus tests for statistical significance meaningless. Skewness and kurtosis describe the shape of a distribution. Skewness measures the extent to which a data set is skewed to one side, for example a long left or right tail on a distribution. A normal distribution has a skewness coefficient of zero, because it is symmetric about the mean. Kurtosis measures the 'thickness' and length of the tails of the distribution. For example a distribution with longer and thicker tails than the normal distribution has an *excess* kurtosis coefficient greater than 0.

70. The distribution of the residuals for model 4 are shown below in Figure 7 and Figure 8.

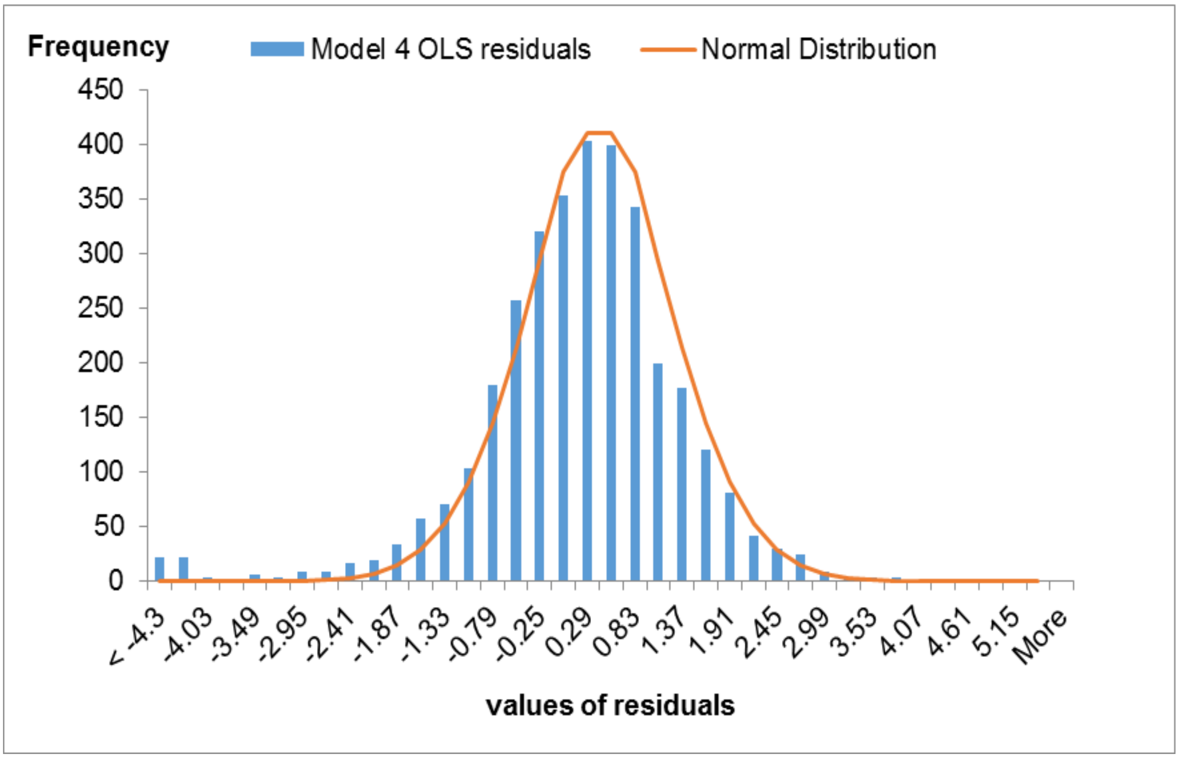
Figure 7 Distribution of Model 4 MM regression residuals based on market corrected data



Source: ERA Analysis

71. The distribution of the residuals from the MM regression of model 4 exhibit a high degree of skewness and excess kurtosis relative to the normal distribution shown as the orange line.

Figure 8 Distribution of Model 4 OLS regression residuals based on market corrected data



Source: ERA Analysis

72. The distribution of the residuals from the OLS regression of model 4 suffers less from skewness and kurtosis. The sensitivity tests for model 4 in Figure 1 do not show this advantage of the OLS regression. However, the residuals still exhibit a high degree of skewness and kurtosis as shown in Table 3.

Table 3 Measures of distribution shape of model residuals

Model	skewness	excess kurtosis
OLS		
1	-2.50	24.90
2	-14.90	463.90
3	-7.70	159.90
4	-3.50	44.20
MM		
1	-2.50	24.60
2	-14.90	464.10
3	-7.70	160.30
4	-4.00	47.15

Source: ERA Analysis

73. Data with a skewness coefficient between -0.5 and 0.5 is approximately symmetric. Data with a skewness coefficient less than -1 or greater than 1 has a highly skewed distribution. A normal distribution has *excess* kurtosis of 0. The results show that model 2 has the highest negative skewness and excess kurtosis. This indicates that the standard errors and thus tests for statistical significance of parameters from model 2 are subject to the greatest uncertainty.

74. The Jarque-Bera test is a formal test of whether skewness and kurtosis differ significantly from that of the normal distribution. Results of the tests are shown below.

Table 4 Jarque-Bera tests of normality in the distribution of regression residuals

Model	JB Statistic	p-value
OLS		
1	89119.14	<0.0001
2	29830067	<0.0001
3	3560467	<0.0001
4	276381.7	<0.0001
MM		
1	87080.25	<0.0001
2	29824957	<0.0001
3	3573903	<0.0001
4	315466.9	<0.0001

Source: ERA Analysis

Note: The MM regression results here are shown only for completeness and should not be attributed any weight. This is because the MM regression residuals are predisposed to failing normality tests.

75. Jarque-Bera tests for normality in the distribution of the residuals reject the hypothesis of normality even at very low levels of significance which allow for much greater deviation from a normal distribution. For the OLS regression, this indicates that the standard errors and thus tests for statistical significance of parameters from all of the models are subject to considerable uncertainty. Further details and analysis of skewness and kurtosis are given in Appendix 2.

76. A model that fits the data well would be expected to produce residuals with a similar *shape* of distribution given low or high values of the predictor variable. The distribution of residuals should not change shape significantly depending on whether the residuals associated with high or low values franking are being observed. That is, there should be no relationship between distribution shape and franking credit values, otherwise a model that captures the relationship is more appropriate.

77. The following examines the residuals for models 1 through to 4 ranking the residuals in ascending order by franking credit value. The residuals are ranked in this order to identify patterns that are dependent on theta in a similar way to Figure 5 and Figure 6.[31] Again, patterns in the residuals indicate that the model inadequately captures dynamics in the data. An illustration of this ranking is shown for the model 1 OLS regression in Figure 9.

[31] Note that Figure 5 and Figure 6 are looking at variance or *dispersion* that is dependent on the level of franking credits. Here the emphasis is on distribution *shape* that is dependent on the level of franking credits.

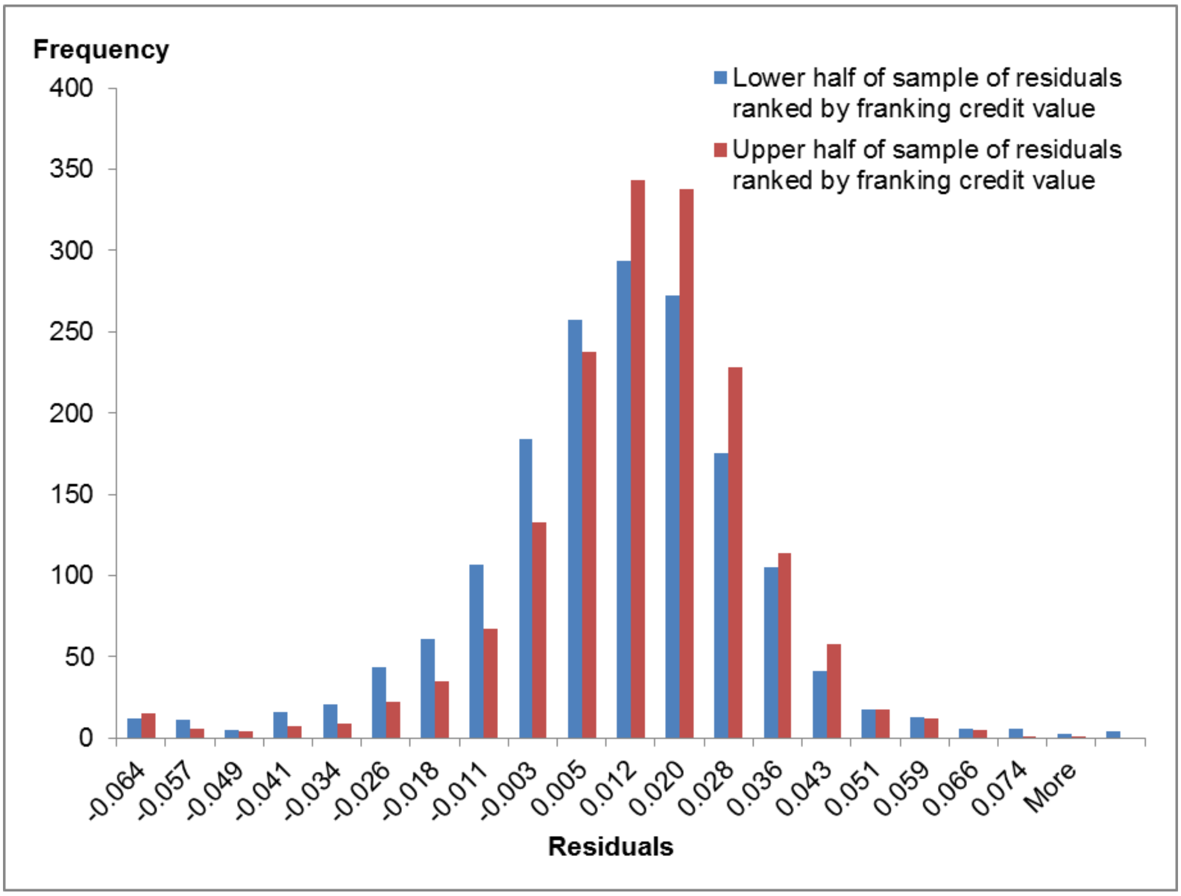
Figure 9 **Model 1 OLS regression residuals plotted against corresponding franking credit values (market corrected data)**

Source: ERA Analysis

Note: For the purposes of clear illustration this figure zooms in excluding less than 1 per cent of the observations.

78. Casual observation of the data points suggest that aside from the variance changing with the value of franking credits (examined in the heteroskedasticity tests), the direction of the skewness in the residuals also changes as the value of franking credits increases. That is, the residuals in Figure 9 are not symmetrically *balanced* around the X axis and the asymmetrical balance from one side to the other appears dependent on the value of franking credits. For values of franking credits close to zero the residuals appear negatively skewed, as the residuals circled in red nearest the lower end of the franking credit value axis are large and negative. For values of franking credits around or greater than 0.5, the residuals appear to be positively skewed: the residuals circled in red nearest the upper end of the franking credit value axis suggest positive skewness. To formally test the changing skewness and shape in the residuals, more generally, the residuals were ranked by their associated franking credit value from lowest to highest. The ranked sample was then split into a lower half and upper half. The distributions on each half of the sample were compared to see if they differ significantly. A graphical illustration of the idea underlying the formal test is shown in Figure 10.

Figure 10 Comparison of residual distribution between lower and upper half of residual sample based on franking credit value - Model 1 OLS regression residuals



Source: ERA Analysis

79. Histograms on each half of the sample of residuals from the OLS regression of model 1 ranked by franking credit value were graphed and compared. Casual observation of Figure 10 indicates that the distributions are considerably different. A formal test of this is the Kolmogorov-Smirnov (K-S) test. This test compares one distribution to another to determine if they are significantly different. The null hypothesis is that the distributions being compared follow each other. A D-statistic larger than the critical value or alternatively a small p-value indicates rejection of this hypothesis.[32] In the current context this test should produce meaningful results because each half of the sample contain 1654 or 1655 observations. The results are shown in Table 5.

[32] The D-statistic is the supremum of the differences between the empirical cumulative distributions of the two samples. The asymptotic distribution of the D-statistic is taken to be the Kolmogorov distribution, for which critical values may be derived.

Table 5 Kolmogorov-Smirnov tests for changes in distribution of residuals

Model	Regression type	D-statisitic	p-value	outcome
1	OLS	0.5385	<0.0001	Reject hypothesis that distributions are the same
	MM	0.5225	<0.0001	Reject hypothesis that distributions are the same
2	OLS	0.4944	<0.0001	Reject hypothesis that distributions are the same
	MM	0.4886	<0.0001	Reject hypothesis that distributions are the same
3	OLS	0.5336	<0.0001	Reject hypothesis that distributions are the same
	MM	0.5258	<0.0001	Reject hypothesis that distributions are the same
4	OLS	0.5896	<0.0001	Reject hypothesis that distributions are the same
	MM	0.5424	<0.0001	Reject hypothesis that distributions are the same

Source: ERA Analysis

80. The K-S tests reject the hypothesis that distributions of the residuals across high and low franking credits follow each other even at the .001 per cent level of significance. This is evident in the reported p-values all being less than 0.0001. The distributions of residuals for all models therefore differ significantly between high and low values of franking credits, in terms of skewness and likely kurtosis. This indicates that the models do not fit the data well, invalidating the resulting parameter estimates. The implication is that a model specification that captures the variable shape and spread of the data will be more appropriate.

81. To summarise, the residuals resulting from the fitting of model 1 through to 4 violate the CLRM assumptions in several ways. A classical linear regression model is not appropriate for any of the data in the models except for model 2. The assumption of variability and random sampling in the predictor variables may be compromised by the data filtering methods and outlier removal processes discussed in section 3.2. Model 4 violates the assumption of no heteroskedasticity in the residuals.

82. All of the four models exhibit considerable skewness and kurtosis in their residuals. This is a symptom of poor fit to the data, given the somewhat extreme behaviour of the data, and compromises the reliability of standard errors and tests for statistical significance of theta. The skewness and kurtosis in the residuals is considerably worse for Model 2, despite it showing no clear violations of the CLRM assumptions. The distribution of residuals from all models differ significantly depending on whether the value of franking credits is low or high. That is, the distribution of the residuals is dependent on at least one of the predictor variables. This is another symptom of poor model validity, and suggests that a specification that accurately captures the dependencies of the residuals on the predictor variables is more appropriate.

5 Better fitting models

83. The point estimates, variation and range in theta estimates are largely driven by the choice of scaling factor (see Appendix 2 difference in estimates of theta). The scaling factor is the variable that the entire model is divided by. Although it may appear that four different model specifications shown in equation (3) through to (6) have been assessed, only *one linear specification* has in fact been trialled. This is because the dependent variable (the variable being explained) is different in each model, due to different scalars having been applied in each model. Hence, each model (1 through

to 4) represents a somewhat different dependent variable. Put another way, four different problems have been presented and only one solution has been proposed for each in terms of a linear model specification.

84. The poor fit of the scaled models and associated model uncertainty is not a trivial issue. The following analysis demonstrates that once the poor fit of the standard linear regression on the scaled SFG models is recognised, a more statistically robust specification can produce higher estimates of theta.

85. It is important to note that the behaviour of prices around ex-dividend days or the 'drop-off' is not well understood. Behaviour in the data that cannot be understood, cannot be controlled for and thus is likely to obscure relationships that the dividend drop-off models attempt to explain.[33] For this reason, the analysis is focused only on the statistical aspects (misspecification) of the dividend drop-off models. The theoretical underpinnings are beyond the scope of this paper.

86. In light of the scaling SFG has applied to the models, an optimal regression formulation is likely to be found outside of the linear formulation implemented thus far for models 1 through to 4. This is because the current set of models largely fail to meet the basic CLRM assumptions. The regression diagnostics implemented here suggest that the standard linear model is a poor fit of the data. On this basis, dividend drop-off studies are subject to considerable *model* uncertainty. Some examples of alternative specifications that resolve some of the issues raised above follow.

87. The 'broken-stick' model attempts to overcome nonlinearity in the data while still using linear regression across segments of the data. This specification allows for two or more straight lines of differing slopes to be fitted across various segments of the data. This controls for nonsensical results exhibited in the extremes of the data by effectively fitting different linear models to those extremes. This is in contrast to removing data that causes such results, or down weighting extreme data using methods such as robust regression.

88. The broken-stick model requires an initial choice of 'tranches' or segments of franking credit values across which data behaviour is thought to differ. The upper bound of each of these tranches are referred to as 'breakpoints'. The model results can be sensitive to the choice of the initial number of breakpoints. An objective criteria is needed to assess the initial choice of breakpoints due to the sensitivity of the model to the choice of breakpoint number.

89. The Akaike Information Criterion (**AIC**) is a commonly applied criteria in selecting the best fitting econometric specification. The explanatory power of an econometric model can be improved by adding irrelevant variables, termed 'overfitting' of a model. The AIC penalises the addition of irrelevant variables while favouring the specification with the greatest explanatory power. The model specification with the lowest AIC is preferred. In the current context, increasing the number of breakpoints results in an increased number of variables in the model. The number of break points are chosen by estimating the broken-stick model repeatedly, increasing the number of

---

[33] For example, Lally reviews literature that identifies components of the dividend drop-off that are not attributed to franking credits and are anomalous. He outlines two specifications of the dividend drop-off model that include and constant term which can potentially control for such anomalies. These specifications produce higher theta estimates than those without constants. See Lally. M, *The Estimated Utilisation Rate For Imputation Credits*, 12 December 2012, pp.13-17.

breakpoints each time, and selecting each estimation that results in a new minimum AIC.

90. The optimal model based on this criteria establishes three segments or 'tranches' of franking credit value across which theta differs significantly for models 1 and model 4. The broken-stick specification applied to model 2 did not produce a model that was significantly different from the original specification for model 2. This is to be expected, because the RESET tests for model specification error shown in Table 1 did not reject the hypothesis that the linear regression for model 2 was correctly specified. The broken-stick application to model 3 data was dismissed on account of producing results for the net dividend coefficient that were not feasible. Appendix 3 gives further details on this analysis. The results for model 1 and model 4 are shown in Table 6 and Table 7 respectively.

Table 6 Model 1 - Broken-stick model theta estimates

Theta	Franking Credit Value Range (*scaled by cum price*)		
	0 to 0.002	0.002 to 0.046	0.046 to 0.061
Estimate	-1.724	0.596	4.879
t-statistic	-1.297	4.806	2.380
Significant at 5 per cent level?	No	Yes	Yes
Proportion of Sample	11.4%	88.1%	0.5%

Source: ERA Analysis

91. As mentioned earlier, the broken-stick model controls for possible nonsensical results exhibited in the extremes of the data by effectively fitting different linear models to those extremes. The broken-stick model, therefore, need only produce one set of sensible theta estimates, preferably representing the majority of the data.

92. The results for model 1 report a feasible and statistically significant theta of 0.596 estimated across 88 per cent of the data.

93. The AIC described above, again, can be used as an objective measure to compare the original specification for model 1 against the broken-stick specification. The broken-stick model reports a lower AIC of -15971.5 than the original model which reports an AIC of -15815.97. The broken-stick model is therefore a better fit based on this criteria.[34]

94. The results for model 4 (shown in Table 7) also report a feasible and statistically significant theta of 0.767 estimated across 80.8 per cent of the data.

95. The reported AIC for model 4 using the broken-stick specification is 10278.24 which is lower than the AIC for the OLS specification for model 4 (10914.09). Again, this

[34] The MM regression of the original specification for model 1 reports an AIC of -15754. This AIC value is higher as the AIC is based on the maximum likelihood. MM regression is not maximum likelihood, and so is not optimal according to the AIC criterion, namely because MM regression does not attempt to fit extreme data well, unlike OLS (the OLS estimator is equivalent to the maximum likelihood estimator under the CLRM assumptions via the Gauss-Markov Theorem).

indicates that the broken-stick model is a better fit. The results for model 4 also report a feasible and statistically significant theta of 0.767 estimated across 80.8 per cent of the data.

Table 7 Model 4 - Broken-stick model theta estimates

Theta	Franking Credit Value Range (*scaled by cum price x volatility*)		
	0 to 0.165	0.165 to 3.715	3.715 to 6.998
Estimate	-1.27	0.767	1.057
t-statistic	-2.068	6.973	1.266
Significant at 5 per cent level?	Yes	Yes	No
Proportion of Sample	18.2%	80.8%	0.1%

Source: ERA Analysis

Note: the ranges differ due to the scaling of franking credit data based on that specified in model 4.

96. It should be noted that the broken-stick model produces estimates that can vary each time the solver software is run.[35] Differing solutions are reflected in the standard errors of theta estimates and ranges. The broken-stick model, however, will tend to produce a better fit. The main point of this analysis is to emphasise that a significantly better fit can be achieved through increasing the flexibility of the regression model, although in this instance greater flexibility has come at the cost of decreased numerical stability in the solution.

97. A 'finite mixture-of-regressions' model was also fitted to the data. This model fits a variety of linear regression models to the data simultaneously. Data points are grouped into 'clusters' based on their distance from each regression line, and on the variance of the residuals associated with the regression for each cluster. The software used in this instance however, only allows an intercept of zero for each regression line.[36] This results in a regression technique that is more or less equivalent to MM regression. The modelling in this instance adds little to the analysis beyond what has been stated for MM regression already. Further details and results of the finite mixture of regressions model can be found in Appendix 3 on alternative models.

98. An interaction term was added to the regression to assess the effect on fit. The interaction between net dividends and franking credits is captured by adding an additional term to the regression model which multiplies the scaled net dividends by the scaled franking credits.[37] A statistically significant coefficient for this interaction term could indicate a model specification error in the form of variable omission. The interaction models reporting sensible results (non-negative theta) and a better fit according to the AIC are reported in Table 8.[38] The results for the remainder of the models are reported in the discussion of the interaction model in Appendix 3.

---

[35] This is due to a stochastic initialisation of parameter estimates.

[36] The finite mixture-of-regressions was implemented using the R package flexmix.

[37] 'Scaled' refers to the fact that model 1 through to 4 divide each variable by a specified factor such as cum price in model 1 or net dividend in model 2.

[38] As indicated by the lower AIC.

Table 8 Interaction model results based on market corrected data

Model	Coefficient	Net Dividend	Theta	Interaction term	Interaction model AIC	Original model AIC
1	**OLS**	**0.835**	**0.624**	**-11.10**	**-15894.02**	**-15815.97**
	t-statistic	-19.9	-5.2	9.0		
	Statistically significant at 5 per cent?	*Yes*	*Yes*	*Yes*		
3	**MM**	**0.856**	**0.047**	**0.004**	**39910.14**	**39916.77**
	t-statistic	-21.4	-0.3	-4.0		
	Statistically significant at 5 per cent?	*Yes*	*No*	*Yes*		
4	**OLS**	**0.867**	**0.720**	**-0.180**	**10473.64**	**10914.09**
	t-statistic	-22.2	-6.9	22.5		
	Statistically significant at 5 per cent?	*Yes*	*Yes*	*Yes*		
	MM[39]	**0.920**	**0.616**	**-0.155**	**10501.47**	**11104.54**
	t-statistic	29.6	7.4	-23.5		
	Statistically significant at 5 per cent?	*Yes*	*Yes*	*Yes*		

Source: ERA Analysis

99. The theta estimates in Table 8 *cannot be interpreted as a stand-alone estimate*. By design, theta in the interaction model is conditional on the level of scaled net dividend.[40] Table 29 in Appendix 3 outlines the value of theta at various quantiles of net dividend. Excluding negative estimates, Table 29 shows that the interaction model produces:

- OLS theta estimates ranging from 0.230 to 0.613 for model 1 data;
- MM theta estimates ranging from 0.068 to 0.543 for model 3 data;
- OLS theta estimates ranging from 0.298 to 0.715 on model 4 data; and
- MM theta estimates ranging from 0.253 to 0.611 on model 4 data.

100. These results confirm that a nonlinear specification is more appropriate for the model 1 and 4 data. The estimates based on these models range from 0.07 to 0.72. However, it should be noted that the estimates at the lower end of this range are sourced from model 3, and hence are subject to higher parameter uncertainty on account of greater skewness and kurtosis in the residuals than models 1 and 4.

[39] Numerical instability was observed in deriving the model 4 estimates using MM regression on the market corrected data. The estimates provided here have a lower AIC than the alternatives to which the MM algorithm converged.

[40] Scaled meaning divided by the variable specified in either model 1, 2, 3 or 4 in section 2.

Table 9 Skewness and kurtosis in residuals of original and additional models

Estimation Method	Model	skewness	excess kurtosis
OLS	**1**	-2.50	24.90
	2	**-14.90**	**463.90**
	3	**-7.70**	**159.90**
	4	-3.50	44.20
MM	**1**	-2.50	24.60
	2	**-14.90**	**464.10**
	3	**-7.70**	**160.30**
	4	-4.00	50.15
Additional specifications	Broken Stick model 1	-2.41	26.34
	Broken Stick model 4	-3.48	56.09
	Interaction model 1 OLS	-2.35	26.06
	Interaction model 3 MM	**-7.70**	**160.26**
	Interaction model 4 OLS	-3.26	51.33
	Interaction model 4 MM	-5.53	68.90

Source: ERA Analysis

101. Table 9 shows that despite achieving a better fit with a range of more flexible models skewness and kurtosis in the residuals remains an issue. Specifications based on model 2 and 3 data exhibit greater skewness and kurtosis. Both of these specifications divide data by net dividends. While this makes the data more amenable to fitting a linear model (see Table 1) the trade-off appears to be very poor model fit in terms of the distribution of the residuals.

102. A high degree of skewness and kurtosis in the residuals in the presence of heteroskedasticity (as seen in the original specification of model 4) is likely to be more of an issue than extreme skewness and kurtosis in the absence of heteroskedasticity, as is the case with model 3. The estimates based on the additional specifications of model 4 are therefore likely to be less robust than the estimates of model 3 if heteroskedasticity has not been accounted for in the model specification and is still present in the residuals. Table 10 shows that heteroskedasticity is no longer present in the broken stick specification for model 4. This provides some assurance that the estimates based on the broken stick specification for model 4 are likely to be more valid than model 3 despite the high degree of skewness and kurtosis in both models.

103. Table 10 shows that heteroskedasticity remains an issue for broken stick model 1, the OLS interaction model 1, and both regressions for interaction model 4. Combined with the high skewness and kurtosis in the residuals of these models, it is likely that these models are a poorer fit, and hence produce estimates with a higher degree of uncertainty than broken stick model 4 and interaction model 3.

Table 10 Breusch-Pagan tests for heteroskedasticity in original and additional models

Estimation Method	Model	BP Statistic	p-value	outcome
OLS	1	3.09	0.08	Do not reject hypothesis of no heteroskedasticity
	2	0.01	0.94	Do not reject hypothesis of no heteroskedasticity
	3	1.06	0.30	Do not reject hypothesis of no heteroskedasticity
	4	60.27	0.00	Reject hypothesis of no heteroskedasticity
MM	1	3.09	0.08	Do not reject hypothesis of no heteroskedasticity
	2	0.01	0.94	Do not reject hypothesis of no heteroskedasticity
	3	1.06	0.30	Do not reject hypothesis of no heteroskedasticity
	4	**60.27**	**0.00**	**Reject hypothesis of no heteroskedasticity**
Additional specifications [41]	**Broken Stick model 1**	**18.59**	**0.01**	**Reject hypothesis of no heteroskedasticity**
	Broken Stick model 4	4.018	0.78	Do not reject hypothesis of no heteroskedasticity
	Interaction model 1 OLS	**17.80**	**<0.001**	**Reject hypothesis of no heteroskedasticity**
	Interaction model 3 MM	1.695	0.429	Do not reject hypothesis of no heteroskedasticity
	Interaction model 4 OLS	**6.17**	**0.046**	**Reject hypothesis of no heteroskedasticity**
	Interaction model 4 MM	**6.17**	**0.046**	**Reject hypothesis of no heteroskedasticity**

Source: ERA Analysis

104. Table 11 summarises all of the models and the outcomes of the diagnostics on their residuals. The models are ranked starting from uncertain to relatively less certain. This is because all of the models are associated with considerable uncertainty on account of the distribution shape being dependent on the residuals and high skewness and kurtosis.

---

[41] The results for the broken stick model are to some degree stochastic in that they vary slightly each time the regression is run. However, repeated estimation does not lead to a different conclusion.

Table 11 Summary of the models and the outcomes of the diagnostics on their residuals ranked from uncertain to relatively more uncertain

Model Specification	Distribution dependent on predictor variables	Skewness and kurtosis	Heteroskedasticity	Fit	Range (2 dp)[42]
Broken Stick model 4	unresolved	high	not present	Best model 4 fit	0.77
MM Interaction model 3	unresolved	higher than model 4, lower than model 2	not present	Best model 3 fit	0.07 - 0.54
Original model 2	unresolved	highest	not present	Better than original model 1, 3 and 4 as linear	0.36
Original model 1	unresolved	high	not present	Highest model 1 AIC	0.21
Original model 3	unresolved	higher than model 4 and 1, lower than model 2	not present	Higher model 3 AIC than interaction model	0.42
Broken Stick model 1	unresolved	high	present	Lowest model 1 AIC, but heteroskedasticity present	0.60
OLS interaction model 1	unresolved	high	present	Lower model 1 AIC than original model, but heteroskedasticity present	0.23 - 0.61
OLS Interaction model 4	unresolved	high	present	Lower model 4 AIC than original model, but heteroskedasticity present	0.30 - 0.72
MM Interaction model 4	unresolved	high	present	Lower model 4 AIC than original model, but heteroskedasticity present	0.25 - 0.61
Original model 4	unresolved	high	present	Highest model 4 AIC	0.33

Source: ERA Analysis

Note: A lower AIC indicates a better fit.

[42] Original model figures based on the average of market corrected OLS and MM theta estimates in Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013. A negative value for theta was excluded from model 4.

105. The nonlinear broken stick models indicated a superior fit for model 4 over the linear model. This model had lower skewness and kurtosis in the residuals than the MM interaction model for model 3. More importantly, unlike the original specification, this model did not exhibit heteroskedasticity in the residuals. The MM interaction model 3 indicated the best fit for model 3. Skewness and kurtosis in the residuals was greater than that of broken stick model 4 and accordingly it is ranked as more uncertain in Table 11. The original specification of model 2 has greater skewness and kurtosis than MM interaction model 3. For this reason it is ranked as more uncertain than MM interaction model 3. The original specification of model 2 ranks above those below it on account of being the only model amenable to a linear specification. The AIC for the original specification of model 1 shows it is a relatively poor fit. However it has lower skewness and kurtosis in the residuals than model 3 and so ranks above it.

106. The models ranked below the original specification of model 3 in Table 11 exhibit heteroskedasticity. This, in combination with the high skewness and kurtosis indicates a very poor fit and greater uncertainty in the estimates. All of the nonlinear models (broken stick and interaction models) in this section of the ranking indicate a better fit than the original specifications. It is difficult to distinguish each of them in terms of ranking, however they all rank above the original model 4 specification. This is because the original model 4 specification exhibits heteroskedasticity in the residuals and indicates a relatively poor fit based on the AIC.

107. The models outlined in this section are not proposed as the ultimate solution to the failings of the original specification identified in section 9. It should be recognised that these models violate the principle of value additivity.[43] That is, receipt of $100 worth of franking credits in two separate instances would not be valued differently to $200 received in one instance. From this perspective, it does not make sense for the value attached to $1 of franking credit to change with quantity of franking credits received. However, the purpose of this model is not to test any such hypothesis. The data exhibits behaviour that is not well understood as demonstrated by the very poor fit of the SFG/Frontier models. The purpose of these models is to demonstrate that accommodating the data (by controlling for this behavior) rather than modifying it, can result in better fitting models and higher estimates of θ.

108. The models outlined here do little to overcome the issue of skewness and kurtosis in the residuals, regardless of which model is fitted. Moreover the skewness and/or kurtosis appears to be dependent on the level of franking credits. This adds to the uncertainty in theta estimates. Models 2 and 3, which divide all variables by net dividends, exhibit much greater skewness and kurtosis in the residuals than models 1 and 4. This indicates that poor model fit and high uncertainty in estimates based on model 2 and 3 variables will remain, even following the application of more sophisticated models.

[43] See Lally. M, *The Estimated Utilisation Rate For Imputation Credits*, 12 December 2012, p.16 for additional specifications that are consistent with the value additivity principle.

Despite this, the broken-stick model estimate that covers the bulk of the data could be interpreted as being the single appropriate value to be used in calculating the utilisation of franking credits.

6 Conclusion

109. SFG and Frontier Economics have raised concerns regarding the sensitivity analysis undertaken in Vo et al. which was relied on by the ERA to inform the range for gamma in its Rate of Return Guidelines. SFG and Frontier Economics partly attribute the Vo et al. finding of instability in estimates of theta to the omission of applying market correction to the data. This study has reproduced the sensitivity analysis in Vo et al. using the same data *with* the market correction applied. The results only minimally support SFG and Frontier's assertions. The market corrected data tends to produce stability plots that fluctuate less than those based on data without the market correction, with one-at-a-time removal of outliers. However, the scale of fluctuation tends to remain large relative to the value of the theta estimate. Estimates of θ also change considerably depending on the model and regression technique chosen.

110. The stability tests also induce a false impression of stability through removing or mitigating the effect of extreme observations using robust regression or removing pairs of observations with offsetting effects. Instead of specifying a model that complies with the data, the focus of the analysis by SFG and Frontier has been to eliminate or mitigate the effect of these data points to create similar looking sets of theta estimates. The similar or 'stable' looking sets of theta estimates produced in the stability plots provide little information about the certainty of parameter estimates or robustness of the estimates to possible violations of basic statistical assumptions. Moreover, substantial instability still exists for the various models' estimates of theta.

111. In applied econometrics it is conventional to diagnose for violations of the CLRM assumptions. This assists in selecting models that produce *meaningful* parameter estimates (such as theta). Such techniques involve visual analysis of the residuals and testing via a suite of statistical procedures. These basic diagnostics have received little attention in Australian theta studies thus far.

112. The application of diagnostics shows that SFG's scaled models violate the CLRM assumptions in the following ways:

- Models 1, 3 and 4 fail the assumption that the linear model is appropriate. In addition, Model 4 violates the assumption of constant variance (or no heteroskedasticity).
- All of the models show a high degree of skewness and kurtosis in the residuals, which violates the assumption that the models' prediction errors are normally distributed.
- In addition to all model residuals exhibiting considerable skewness and kurtosis, the shape of the distribution of residuals has been observed to change significantly with the franking credit value. This indicates that the models do not adequately capture important relationships in the data.
- The presence of heteroskedasticity in combination with high skewness and kurtosis in model 4 is likely to produce particularly unreliable estimates.

113. SFG expressed a preference for model 4 on the basis of greater stability and low standard errors. Model 4's violation of the assumption of constant variance (or no heteroskedasticity) indicates that the small standard errors are unreliable and that

model 4 is not a wholly valid specification.[44] The sensitivity analyses applied by SFG, Frontier and Vo et al. fail to diagnose this violation.

114. It is important to note that the behaviour of prices around ex-dividend days or the 'drop-off' is not well understood. Research shows that other factors may be present in dividend drop-off data that are not well understood or controlled for.[45] These factors may obscure the relationship between franking credits and the drop-off in the DDO models resulting in model misspecification, poor fit and model uncertainty. In light of this uncertainty and the poor fit of the SFG models, this study has trialled other regression formulations. The theoretical underpinnings explaining observed relationships are beyond the scope of this paper.

115. A nonlinear 'broken stick' model based on model 4 data resolves the issue of heteroskedasticity in the residuals and produces a fit superior to the original specification. The broken stick model also has lower skewness and kurtosis than model 2 and 3. On this basis, broken stick model 4, although still producing very uncertain estimates, produces more certain estimates than the other specifications in this study. The diagnostics show that the other specifications in this study are less robust than this model and produce estimates ranging from 0.07 to 0.72. Model 4 shows that much higher theta estimates than those proposed by SFG/Frontier are possible once the issues of model fit are taken into account.[46]

116. All models in this study do not overcome the issue of skewness and kurtosis in the residuals and the dependency of distribution shape on the predictor variables. This finding renders theta estimates from the dividend drop-off method highly uncertain.

---

[44] Professor Lally also expresses concerns over SFG's preference for model 4. See Lally. M, *The Estimated Utilisation Rate For Imputation Credits*, 12 December 2012, p.4.

[45] Lally. M, *The Estimated Utilisation Rate For Imputation Credits*, 12 December 2012, pp.13-17.

[46] Specifically model 4 produces a point estimate of 0.77 for theta across 80 per cent of the data.

Appendix 1 Models for the Estimation of θ

1. The notation applied in this appendix mirrors that of Vo et al., but includes corrections for typographical errors and any confounding use of notation (e.g., where an index or symbol was used to describe more than one list of objects).

2. The *dividend drop-off* (DDO) may be estimated as the price drop-off from the cum-dividend day price ($P_{c,i}$) to the ex-dividend day price ($P_{x,i}$) for dividend event i. The expected price drop-off is therefore defined as a proportion of the face-value the net dividend and the franking credits that an investor in the asset is entitled to:

$$E\left[P_{c,i} - P_{x,i}\right] = \delta D_i + \theta FC_i$$

Where:

D_i is the face value of the net dividend;

FC_i is the face value of the franking credit; and,

δ and θ are the extent to which the net dividend and the franking credit are valued by the investor, respectively. The terms δ and θ are expected to be less than 1, indicating that investors value the net dividend and franking credit at less than the face value.[47]

3. The choice of rescaling factor varies with each of the four statistical models considered by SFG and the ERA.[48] These models are specified in Table 12 , but have the general form:

$$\frac{DDO_i}{S} = \frac{\widehat{\delta} D_i}{S} + \frac{\widehat{\theta} FC_i}{S} + \varepsilon_i$$

Where:

$DDO_i = P_{c,i} - P_{x,i}$ is the observed (unadjusted) dividend drop-off at the time of the i^{th} dividend event;

$\widehat{\delta}$ and $\widehat{\theta}$ are estimates of δ and θ given scaling factor S ;

ε_i is a (normally distributed) random error term given by $\varepsilon_i \sim N\left(0, \sigma^2\right)$; and

σ^2 is the variance of the residual error term to be estimated in the model.

[47] As noted above, issues relating to whether dividend drop-off studies correctly estimate theta as defined are not examined in this statistical review. The ERA has observed that dividend drop-off studies may not correctly estimate the utilisation rate (Economic Regulation Authority, *Final Decision on Proposed Revisions to the Access Arrangement for the Dampier to Bunbury Natural Gas Pipeline 2016 – 2020*, Appendix 5 – Gamma, 30 June 2016, pp. 11-12).

[48] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7, 21 March 2011,* paragraph 12, p. 62.
Estimating the Market Value of Franking Credits: empirical Evidence from Australia', *Beyond the Frontiers: New Directions in Economics Held at Murdoch university, Perth, Western Australia on 7-10 July, 2013, Proceedings of the 42nd Australian Conference of Economists,* Perth, Western Australia, Murdoch University, 2013, Table 1.

4. Alternatively, DDO_i may be replaced with a market corrected value, namely $DDO_i^* = P_{c,i} - P_{x,i}^*$, where the ex-dividend price is adjusted by the daily market return $r_{x,i}$ on the ex-dividend date:

$$P_{x,i}^* = P_{x,i} \big/ \left(1 + r_{x,i}\right)$$

5. The standard error $\hat{\sigma}_i$, of the mean return associated with an asset for which a dividend event was observed, may be included in the scaling factor S. This standard error may be defined as:

$$\hat{\sigma}_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\omega_{i,\mathrm{t}(i)-5-j} - \bar{\omega}_{i,t(i)}\right)}$$

Where:

$r_{i,\mathrm{t}(i)}$ is the individual asset return at time $t(i)$ associated with dividend event i;

$r_{i,\mathrm{t}(i)}^m$ is the market return at time $t(i)$ associated with dividend event i;

N is the number of market trading days over the past calendar year;

$\omega_{i,\mathrm{t}(i)} = r_{i,\mathrm{t}(i)} - r_{i,\mathrm{t}(i)}^m$ is the excess asset return above the market return at time $t(i)$ associated with dividend event i; and

$\bar{\omega}_{i,t(i)} = \frac{1}{N}\sum_{j=1}^{N}\omega_{i,\mathrm{t}(i)-5-j}$ is the mean excess asset return over the past year, excluding returns within five days prior to a dividend event.

Table 12 Parametric form of DDO equations used by SFG for estimation of theta.[49]

Model	Parametric Form[50]	S
1	$\frac{P_{c,i}-P^*_{x,i}}{D_i}=\hat{\delta}^{(1)}+\frac{\hat{\theta}^{(1)}FC_i}{D_i}+\varepsilon_i^{(1)}$	D_i
2	$\frac{P_{c,i}-P^*_{x,i}}{P_{c,i}}=\frac{\hat{\delta}^{(2)}D_i}{P_{c,i}}+\frac{\hat{\theta}^{(2)}FC_i}{P_{c,i}}+\varepsilon_i^{(2)}$	$P_{c,i}$
3	$\frac{P_{c,i}-P^*_{x,i}}{D_i\hat{\sigma}_i}=\frac{\hat{\delta}^{(3)}}{\hat{\sigma}_i}+\frac{\hat{\theta}^{(3)}FC_i}{D_i\hat{\sigma}_i}+\varepsilon_i^{(3)}$	$D_i\hat{\sigma}_i$
4	$\frac{P_{c,i}-P^*_{x,i}}{P_{c,i}\hat{\sigma}_i}=\frac{\hat{\delta}^{(4)}D_i}{P_{c,i}\hat{\sigma}_i}+\frac{\hat{\theta}^{(4)}FC_i}{P_{c,i}\hat{\sigma}_i}+\varepsilon_i^{(4)}$	$P_{c,i}\hat{\sigma}_i$

Source: ERA Analysis

6. In principal, $\hat{\sigma}_i$ should be estimated alongside the model parameters within a generalised least squares (GLS) framework. In practice, $\hat{\sigma}_i$ forms a *plug-in* estimator that is treated as known (and without uncertainty) within an ordinary least squares (OLS) framework.

7. In summary, the key decision variables related to the choice of model being applied to the data are:

- The scaling factor (e.g., dividends or cum-dividend price)
- The estimation method (e.g., OLS, MM, or LAD)
- The market correction factor (to include or exclude the market correction factor).

8. The data applied in analysis of the above models in following sections is the same as those data prepared by Vo et al.[51] Hence, the data are comprised of dividend events between 1 July 2001 and 30[t] June 2012.

[49] Note that the order of the models, as described by the different scaling factors, is labelled differently from that reported in Table 3 of Vo et al. (2013). SFG (2011) estimates conflict between their Figure 9, and what is reported in their Table 5 and Table 6. Here, we follow the order of models reported by Vo et al. (2013).

SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7, 21 March 2011,* paragraph 56, p. 21.

Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013.

[50] The notation $\bullet^{(k)}$ is a quantity or estimate relating to the k^{th} model specifically.

[51] Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013. Section (v) of the analysis.

Appendix 2 Regression Diagnostics

Potential Issues in Regression Modelling

1. This section describes, in brief, different features of the data that may impact on the accuracy of a regression model, namely: outliers, multicollinearity, nonlinearity and heteroskedasticity. These potential issues are firstly defined, their method of identification described, and their possible resolution outlined. These descriptions provide the conceptual foundation for arguments that will be put forward in later sections to evaluate criticisms of the ERA's considerations in the estimation of θ that have been put forward by SFG Consulting and Frontier Economics.

Outliers

2. An *outlier* is any data point that is distant from the remainder of the data. In general, outliers may exist because:

   - They occur randomly, particularly when the underlying distribution is heavy-tailed (assigns proportionately more probability to extreme values than does the standard normal distribution).
   - They are the result of some measurement error.
   - They are the result of two or more different data generating processes.

3. If an outlier is the result of a heavy-tailed distribution, either towards one side of the distribution (*skewed*) or both sides (*leptokurtic*) of a measure of the distribution's central tendency, then either modelling the data process as a heavy-tailed distribution or as a nonlinear process is to be preferred. Heavy-tailed behaviour may also be a result of heteroskedasticity.

4. If an outlier is the result of a measurement error then the validity of that data point as representative of the data generating process is likely void. In this case, when measurement error is known to occur, then removal of the outlier is appropriate. Alternatively, a method of model estimation that is robust to outliers may be preferred to actually removing the outliers.

5. If an outlier is the result of two or more data generating processes then the appropriate action is to model the data as two or more processes. This requires a different form of model to be applied that explicitly allows more than two data processes to be modelled (the data captures a multi-modal process).

6. Diagnosis of why different outlier behaviour occurs is relatively straightforward, and includes: [52]

   - Q-Q plots of the regression residuals for deviations from normality or some other distribution assumption, in the case of heavy-tailed distributions.

---

[52] R.D. Cook and S. Weisberg, *Residuals and Influence in Regression*, New York, Chapman and Hall, 1982.

D.A. Belsley, E. Kuh and R.E. Welsch, *Regression Diagnostics: Identifying influential data and sources of collinearity,* Volume 571, Hoboken, New Jersey, John Wiley & Sons, 2005.

In total 39 diagnostic methods were listed in Table 2 to identify outliers by Aguinis et al. (2013).

H. Aguinis, R.K. Gottfredson and H. Joo 'Best-Practice Recommendations for Defining, Identifying, and Handling Outliers', *Organisational Research Methods*, vol. 16, 2013, pp 270-301.

Combined with Monte Carlo simulation Q-Q plots can help test and diagnose whether distributional assumptions are correct.

- Identification of outliers with large influence on the regression. Metrics include Cook's Distance, DFFITs and DFBETAs. The impact of each outlier on the regression can be quantified.
- Added variable and/or residual plots, which include 'component plus residual' plots (**CPR**), that help identify whether nonlinearities or unexplained phenomena still exist within the data after key variables in the regression have been taken into account. These plots have, as axes, the different components of the regression, including residuals, studentised-residuals, fitted-values based on one or few predictors, and the predictor terms themselves (in this case the face value of net dividends and franking credits). Confidence bounds associated with added variable and residual plots can also help diagnose heteroskedasticity.

Multicollinearity

7. *Multicollinearity* occurs when two or more predictor variables in a regression are found to be highly correlated with each other. In general, multicollinearity does not interfere with the overall fit of the model or impair accurate prediction.

8. Multicollinearity does reduce the identifiability of individual predictor terms within a model. This reduced identifiability has two potential impacts. The first is that the standard error associated with each parameter estimate will be higher than if multicollinearity is not present, hence indicating greater uncertainty around parameter estimates. The second impact is that the uncertainty reflects that parameter estimates will be sensitive to changes in model structure or reparameterisations of the data. As such identifying the correct model structure can be difficult when a model needs to be selected from among a range of candidate models, including from different predictor variables.

9. Multicollinearity in the dividend data has previously been noted as a potential issue.[53] As the DDO models are focused on returning an accurate estimate of θ then any multicollinearity in the data will likely result in those estimates being relatively uncertain.

10. Multicollinearity may be identified through calculation of a *variance inflation factor* (VIF). In principle, the VIF measures how much the estimated variance of parameter estimate increases because of collinearity. As a rule of thumb a VIF > 10 indicates strong multicollinearity, and may be interpreted as the standard error of a parameter estimate being more than three times larger than what it would be if the predictor variables were uncorrelated.[54]

11. Ridge regression (including LASSO and elastic net regression), which penalises the size of the regression coefficients, will to an extent counter multicollinearity, resulting

---

[53] Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013.

[54] M.H. Kutner, C.J. Nachtsheim and J. Neter, *Applied Linear Regression Models*, 4th edn., Chicago, McGraw-Hill/ Irwin, 2004. The authors refer to literature on rules-of-thumb for the VIF factor back to Marquardt (1970).

D.W. Marquardt, 'Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation', *Technometrics*, vol. 12. 1970, pp. 591-612.

in mean square error (**MSE**) values that are less than the OLS estimates. However, ridge regression penalises the least important predictors the most, and so the coefficient on franking credits in the dividend drop-off framework will be smaller as a result. This reduction in θ estimates simply results in a better predicting model, rather than return a more accurate estimate of θ. Therefore, ridge regression (or other form of elastic net model) should not be applied in DDO studies.

12. There is no real solution to resolving multicollinearity, especially as there is no scope within the θ estimation problem to select different predictor variables. If multicollinearity exists then the general strategy is to accept that any parameter estimates associated with multicollinearity will have higher standard errors than parameter estimates in the absence of multicollinearity.

Nonlinearities

13. Nonlinearities may result from either a nonlinear combination of the predictor variables, or from multi-modal behaviour. That is, there are separate clusters of data whose regression estimates differ.

14. Generally, nonlinearities can be identified from added variable plots or CPR plots. If a nonlinear process is modelled as a linear process then model residuals may appear to be skewed over all or part of the domain of possible predictor values. Hence model misspecification may at times be confounded with an outlier issue.

15. Once potential nonlinearities are identified then nonlinear regression methods may be explored. The Akaike Information Criterion (**AIC**), or some other measure of predictive accuracy, may be applied to compare linear and nonlinear models to asses which model provides a better fit to the data. Hence, the superiority of nonlinear models to linear equivalents must be demonstrated before the nonlinear model is adopted.

16. A range of possible nonlinear models may be applied. Smooth nonlinear models, such as those supported by basis splines or kernels, are problematic insofar as they derive θ estimates that depend on the value of the predictor variables. In contrast, nonlinear models that can potentially provide a single estimate for the majority of the data include segmented (or broken-stick) models, and finite mixture-of-regression models.

17. In a sense, a nonlinear model results in estimates of θ that are dependent on the values of the predictor variables. This suggests that dependency among the predictor variables should also be considered. What results in this case is an estimate of θ that is explicitly dependent on the value of the net dividend. Such models are readily considered by including an interaction term within the regression model between the franking credit and net dividend.

18. The broken-stick, mixture-of-regressions and interaction models are discussed further in Appendix 3.

Heteroskedasticity

19. Heteroskedasticity exists when variability in the data varies with the data values themselves. The key impact of heteroskedasticity is to invalidate statistical hypothesis tests. For example, the variance of the data can be underestimated when heteroskedastiticty exists and models are being estimated through OLS, but without

necessarily affecting the bias (or not) of the parameter estimate. Moreover, heteroskedastic processes may generate data points that are of high influence on the parameter estimates (generate significant outliers).

20. There are numerous ways of detecting heteroskedasticity. The first is through interpreting a changing spread of data points about a mean trend in regression plots of residuals on predictor variables. Formal tests may follow, with SFG Consulting applying White corrected standard errors in addition to the scaling factors.[55] Thirdly, heteroskedastic variables can be modelled explicitly, and be formally tested for inclusion or exclusion within the regression model.

21. If heteroskedasticity is detected then there are generally two key ways to deal with it. The first is to transform the data to a scale where heteroskedasticity is minimised. This is the idea behind applying the scaling factor S to the dividend drop-off DDO_i outlined above. The second is to apply a model that explicitly accounts for the heteroskedasticity through the inclusion of further scalar parameters. For example, modelling variance as linearly dependent on the value of the franking credit.

SFG and Frontier's Sensitivity Analysis

22. As highlighted above the outlier issue may be confounded with possible nonlinearities and heteroskedasticity in the data. Here, we start with a critique of SFG Consulting's outlier-based sensitivity analysis. In later sections, we then proceed to explore possible nonlinearities in the data, and assess the importance of nonlinearities in the data relative to outliers.

DFBETAs

23. SFG Consulting were first to define a stability analysis of estimates of θ:

> The ex-ante screening and checking of data required by the Terms of Reference is designed to eliminate outlier data points that are erroneous in some respect and which are likely to have had a disproportionate influence on the estimate of theta. Even after having performed this screening and checking process, it is inevitable that some of the remaining data points will be more influential than others. Consequently, we have quantified the sensitivity of our estimates of theta to influential observations by conducting a stability analysis.[56]

24. This stability analysis can be stated simply as the sensitivity of estimates of θ to the sequential removal of the most *influential* observations from the data sets. Statistically, an *influential* observation is one that results in a large change in a parameter estimate or prediction when that observation is omitted from the data set. Common measures of influence in the statistical literature include Cook's Distance, DFFITs and DFBETAs, all of which are conceptually similar. Each of these metrics may be used to identify the most influential points in the data set.

---

[55] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* March 2011, p. 26.

H. White, 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity', *Econometrica*, vol. 48, 1980, pp. 817-838.

[56] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* March 2011, p. 28.

25. Both SFG Consulting and the Vo et al. base their stability analyses on the calculation of the DFBETA statistic, which measures to what extent an observation affects the estimate of a regression coefficient: [57]

$$DFBETAS_{k,p} = \frac{\hat{\beta}_p - \hat{\beta}_p^{(-k)}}{SE\left(\hat{\beta}_p^{(-k)}\right)}$$

Where:

$\hat{\beta}_p$ is the estimate of the p^{th} parameter in the model (in this case $\hat{\theta}$); and

$\hat{\beta}_p^{(-k)}$ is the estimate $\hat{\beta}_p$, but with the k^{th} observation removed.

26. From the text of various SFG reports it appears that only the unstandardised measure $\hat{\beta}_p - \hat{\beta}_p^{(-k)}$ has been quantified, and not the standardized DFBETAs statistic:

> We do this by first determining which single observation, if removed, would result in the greatest increase in our estimate of theta. We then determine which single observation, if removed, would result in the greatest decrease in our estimate of theta. We then remove both observations and re-estimate theta. We then repeat this process by removing another pair of observations. We continue in this manner, removing pairs of observations, until 25 pairs have been removed.[58]

27. A set of DFBETAs values are then derived in a recursive fashion: for the ERA, once the observation with the most extreme absolute value for DFBETAs is identified then this observation is removed from the data set, and the DFBETAs statistic is recomputed for the reduced data set.

28. In Vo et al. this recursion is conducted 30 times, with data removed one-at-a-time where only the most extreme data point is removed at each step of the recursion.[59] For SFG Consulting the recursion was conducted 25 times in 2011, and observations are removed pairwise at each step of the recursion.[60] In 2014 the SFG analysis applies 20 iterations of the recursion with the pairwise removal of points.[61] Pairwise removal implies that the observations with the highest and lowest DFBETAs values are removed at each step of the recursion.

29. SFG Consulting initially claim that for Model 4 ($S = P_{c,i}\hat{\sigma}_i$):

> The stability analysis for Model 4, in Figure 8 above, shows that the estimates of the value of cash dividends, the value of theta, and the value of the combined package

---

[57] Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013. Equation 30.

[58] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, p. 28.

[59] Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013. See Part (i) of Section (VI) Sensitivity Analysis.

[60] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7, 21 March 2011,* p. 28, paragraph 79.

[61] SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21st May 2014, pp. 96-97, Figures 13-16.

are very stable and robust to the removal of pairs of influential data points. That is, the estimates from Model Specification 4 are less sensitive to the effects of influential observations.[62]

30. For the other models with different scaling factors (Models 1-3) then SFG observe that the estimate of θ decreases as more observations are removed.[63] The scale of this decrease in $\hat{\theta}$ varies with the scaling factor, from approximately 0.05 to 0.10. The scale of variability (the range of the 95% confidence band provided) varies approximately from 0.25 (Figure 8) to 0.6 (Figure 5).[64]

31. Vo et al. observes with its method that there is very much a 'see-saw' response in $\hat{\theta}$ to the sequential removal of outliers. The scale of the see-saw pattern ranges from approximately 0.25 to 0.45 for Vo et al.'s market-corrected data, depending on the estimation method being applied (these numbers are for MM estimators, less for LAD estimators, more for OLS estimates).[65]

32. This difference in stability analyses has led SFG to later claim that:

> Vo et al. implement a stability analysis known as the DFBETAs approach. This approach differs from the SFG stability analysis in two primary ways:
>
> a) Influential observations are removed one-at-a-time, rather than in pairs; and
>
> b) The stability analysis is only applied in relation to the non-standard approach whereby prices are not corrected for market movements over the ex-dividend day.[66]

33. The lack of a denominator in the calculation of DFBETAs could be a further point of difference in the stability analyses conducted by SFG Consulting and Vo et al. Calculating a standardized version of an influence statistic, through inclusion of a denominator estimating the variance of the parameters (as in DFBETAs) or of the predictions (as in Cook's distance or DFFITs), is standard practice in outlier detection.[67]

Critique of the DFBETAs Stability Analysis

34. The first point to note is that an extensive number of ex-ante filters have been applied to the data, by both SFG and the ERA, to remove potentially contentious data points from the data set that are either erroneous or irrelevant to the analysis.[68] It is therefore reasonable to conclude in this instance that any outliers that arise from the modelling

---

[62] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* March 2011, p. 28.

[63] Ibid, pp. 29-31.

[64] Ibid.

[65] Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, Appendix 1.

[66] SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21 May 2014, pp. 95, paragraph 462.

[67] R.D. Cook and S. Weisberg, *Residuals and Influence in Regression*, New York, Chapman and Hall, 1982.

[68] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* March 2011, p. 4.

Vo, D., Gellard, B., Mero, S., '*Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013. See part (i) of Section (IV) Analysis.

process are not the result of measurement error. Consequently, a valid analytical perspective is that the removal of data should not be attempted, as it is highly likely that outliers instead arise from either a heavy-tailed spread of the data, or from nonlinear or multi-modal data. For similar reasons, modelling to incorporate multi-modality and/or heavy-tailed phenomena is likely to be preferred to robust methods of estimation or outlier deletion.

35. The logic of applying a recursive set of DFBETAs estimates when identifying outliers must also be questioned for the following reasons:

- The first reason is one of parsimony; whereby DFBETAs estimates derived from a leave-one-out procedure are equivalent to the recursion, but requires much less effort to compute.[69]
- The second reason is that the apparent stability of the SFG outputs is a function of the leverage and influence of the data points. In this instance, one can develop a model with a highly biased estimate of θ, yet produce a stability analysis as constructed by SFG which would lead one to conclude that the model is stable. As this is the only regression diagnostic proposed it does not account for other possible weaknesses in the common linear modelling assumptions.[70]
- The third reason is that SFG do not cite any literature that suggests that their recursive approach to outlier detection and resolution, via their stability analysis, is the most appropriate from the range of possible tools and options that have thus far been developed in the statistical literature for model validation.[71]

36. To highlight the fact that a simpler procedure achieves the same outcomes as the more computationally intensive stability analysis proposed by SFG, consider this example:

- Apply both the SFG procedure to identify the 30 most extreme outliers, and the standard leave-one-out DFBETAs procedure.[72]
- Apply the method across the four scaling factors for OLS estimated models, with or without market correction, to assess the consistency of the results.
- There is 95% overlap between the influential outliers identified through each procedure on average across all of the scaling factor and market correction scenarios (Table 13).

---

[69] The leave-one-out-procedure is differentiated from SFG's recursion process in that it only leaves an observation out for the purposes of determining its leverage. Unlike SFG's recursion process, theta is not recalculated each time an observation is left out.

[70] SFG also propose a number of other stability analyses, but these all involve trimming the most extreme values (apart from applying robust estimation which itself is designed to be applied to data with a preponderance of outliers, and which is discussed elsewhere in this report), and hence must be identified as testing the same assumption of undue outlier influence as what SFG's proposed stability analysis is intended to achieve. See SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, p. 31.

[71] H. Aguinis, R.K. Gottfredson and H. Joo 'Best-Practice Recommendations for Defining, Identifying, and Handling Outliers', *Organisational Research Methods*, vol. 16, 2013, pp 270-301.

[72] R.D. Cook and S. Weisberg, *Residuals and Influence in Regression*, New York, Chapman and Hall, 1982.

37. Hence, the stability analysis of SFG is largely redundant, as a simpler and more commonly applied procedure achieves a very similar outcome in identifying the 30 most extreme outliers. It then follows that the SFG stability analysis can largely be reproduced through a plot of DFBETAs on the outliers, in order of decreasing influence.

Table 13 Percentage overlap in outliers: SFG versus standard DFBETAs procedure

Model	Without Market Correction	With Market Correction
1	96.7	93.3
2	90.0	100.0
3	100.0	100.0
4	93.3	80.0

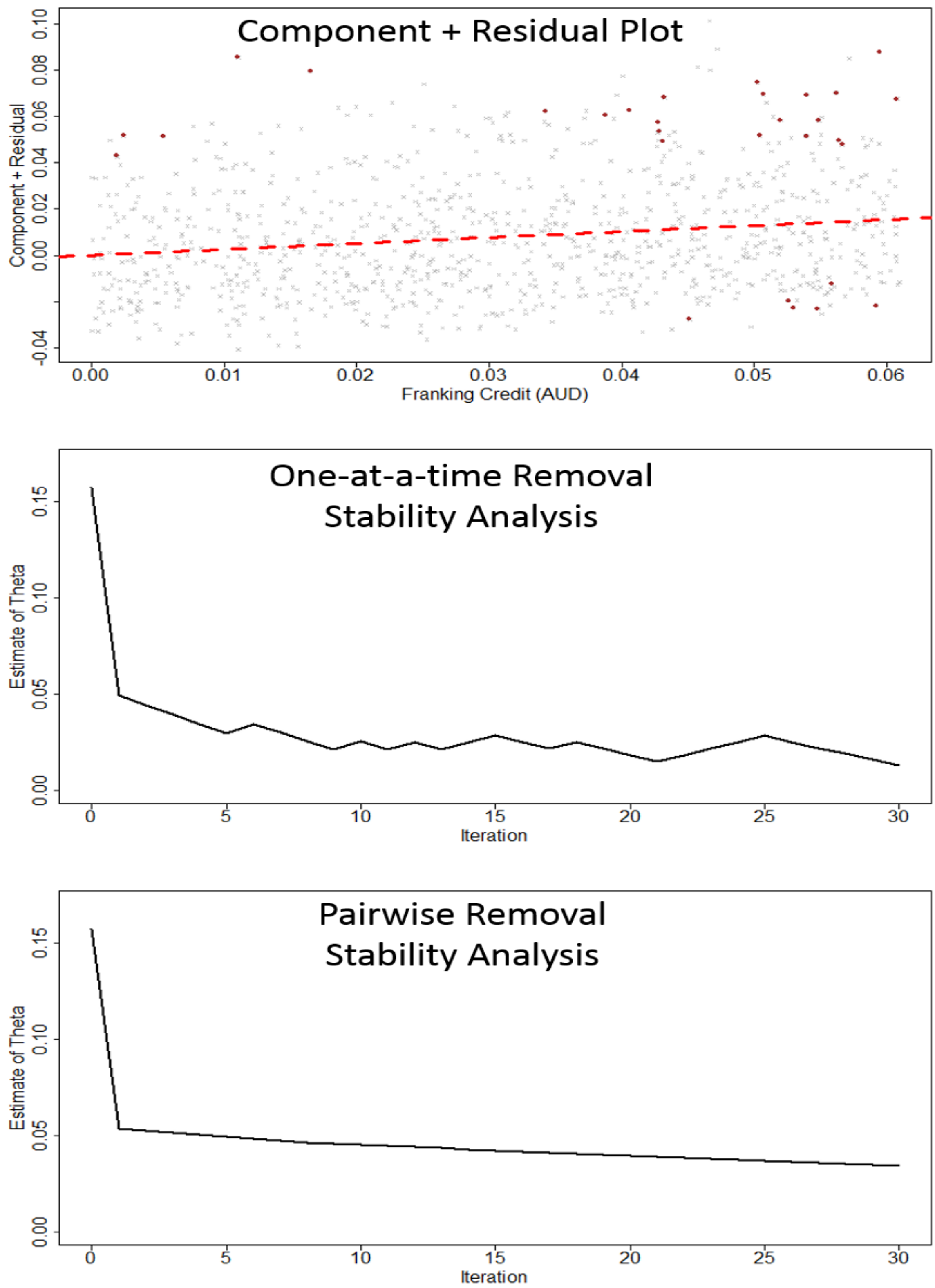
38. The apparent stability of the SFG stability plots is largely predictable if the skewness and heteroskedasticity of the data are known. To demonstrate this fact example data were simulated by applying an error distribution that was positively skewed (Figure 11; top).[73] The outliers were then sequentially removed one-at-a-time (Figure 11; middle) or removed pairwise (Figure 11; bottom). The one-at-a-time removal results in a 'see-saw' response, whereas the pairwise removal results in a smoothed response. Both stability analyses result in a θ estimate that decreases with the sequential removal of outliers, as a result of the positive skewness. These results mimic the difference in smoothness between the ERA's and SFG's stability analyses.

[73] The skewed distribution is generated from normally distributed data following methods in:

C. Fernández and M.F.J. Steel, 'On Bayesian Modeling of Fat Tails and Skewness', *Journal of the American Statistical Association*, vol. 93, 1998, pp. 359- 371.

For simplicity, the simulated data were distributed initially as $DDO_i \sim N\left(0.8D_i + 0.2FC_i; \hat{\sigma}^2\right)$; with skewness parameter $\xi = 2$; $\hat{\sigma}^2$ the plug-in estimate of the residual variance from Model 1, DDO_i equivalent to Model 1 values; and simulation size $M = 1000$.

Figure 11 Stability analyses on positively skewed simulated data



Source: ERA Analysis

39. The 'see-saw' behaviour of the one-at-a-time procedure in this example may be explained by the distribution of the most influential outliers in the CPR plot (Figure 11; top). When the residuals of the regression are positively skewed then the majority of outliers lie above the regression line. Moreover, as the regression has no intercept term then the majority of outliers are also on the right side of the graph, given there is little opportunity to influence the regression from the left side. If one of the large positive residuals on the right side are removed then the slope of the regression will consequently fall. Hence the overall trend is decreasing. However, on occasion one of the large negative residuals on the right side, or large positive residuals on the left side will then be deemed to be the most extreme outlier, resulting in an increase in slope with its removal.

40. The 'smooth' behaviour of the pairwise removal procedure is also self-evident. As the most extreme negative outlier is paired with the most extreme positive outlier, and as the positive outlier is more extreme (has more influence) than the negative outlier, then the overall change in the regression is a decrease in slope. As the data are positively skewed then there are more positive outliers than negative outliers. Pairwise removal therefore leads to a mostly monotonic decrease in the regression slope, and hence the estimate of θ declines in a more or less smooth fashion (depending on the arrangement of extreme values).[74]

41. Qualitatively, there is little difference between the one-at-a-time and pairwise removal procedures in representing skewness in this example. That is, both show an overall decreasing trend in θ with outlier removal in Figure 11.

42. If the data were negatively skewed, relative to the fitted model, then the opposite trend would happen. The stability analysis would illustrate an increase in $\hat{\theta}$ with the sequential removal of outliers. If there is little change in the stability analysis trend then it may reasonably be assumed that the data are more or less distributed symmetrically about the fitted model, and more in keeping with the assumptions of multiple linear regression modelling. Hence, the behaviour of the stability analysis is largely predictable if the skewness (and indeed nonlinearity and heteroskedasticity) of the data is known.

43. Critically, in this simulated example there are no outliers per se. The extreme values are, in probability, expected to occur given the underlying distribution. One may therefore argue that the stability analysis is nothing more than a form of graphical representation of the skewness (or lack of skewness) of the data, at least in this example. There do exist better graphical representations and measures of these features in the statistical and econometric literature (Q-Q plot; Pearson's moment coefficient).

44. That a stability analysis converges to a lower value more likely means that the original model is a poor fit to the data, insofar as it does not satisfy the standard assumptions of a multiple linear regression. Such a non-constant trend in response to outlier removal invalidates the fitted model, at least in part.

---

[74] As stated in paragraph 81 of SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* March 2011, p. 29:

> Figure 5 shows that the original point estimate of theta from Model 1 was 0.16. When the first pair of observations (one observation that would maximally increase the estimate of theta and one that would maximally decrease the estimate of theta) is removed, the point estimate of theta falls to 0.14. As further pairs of observations are removed, the point estimate of theta falls more marginally before levelling off at approximately 0.07.

45. Even when a stability analysis is seen to be 'stable', there still is no guarantee that the model is a good fit of the data. Nonlinearities in the data may still result in a distribution of influential observations relative to the model that 'cancel' each other out, and produce an overall flat response to outlier removal. Overall, one should conclude that little meaningful information can be inferred from the stability analysis beyond common regression diagnostics.

Robust Models

46. In addition to the stability analysis SFG provide a set of sensitivity analyses and robustness checks, as proposed in part by the AER.[75] The application of a robust model (the MM estimator) is given the most weight from among these methods (including from among three other candidate robust methods).

47. There is a logical contradiction when simultaneously applying robust methods while removing outliers to the estimation of θ. This is because robust methods are designed specifically to largely avoid (give lower weight to) extreme values. Hence, it is expected that outliers under robust methods will have little influence on regression estimates. Consequently, any stability analysis applied to a robust method of the type applied by SFG or Frontier Economics will show only that the robust method is stable. In the context of robust analysis a DFBETAs stability analysis as proposed by SFG is therefore largely meaningless.

48. Robust methods safeguard model estimates from observations that are either a notionally unreasonable value, or that have an undue influence on the regression estimates. The unreasonableness of highly influential points may be the result of heteroskedasticity, although with large sample sizes heteroskedasticity can result in an unbiased estimate of model parameters. Robust methods do not, however, safeguard against nonlinearities in the data.

49. To demonstrate this the CPR plots of OLS and MM regressions may be applied to the data (Figure 12 and Figure 13). The majority of data are observed as having a franking credit value (adjusted for the cum-dividend date price as per model 1) of between 0 and 0.02, with 10.8% of the entire data set having a franking credit of 0 AUD associated with each dividend event. The majority of the net dividend data occur at a scaled value of less than 0.06.

50. The remainder of the observations in Figure 12 and Figure 13 (those scaled net dividend observations greater than 0.06 and scaled franking credit observations greater than 0.02) may be seen to be highly influential as the nonlinear fit (via a lowess procedure) departs from the linear trend at this point.[76] The nonlinear trend line decreases following the negative outliers on the right side of both the franking credit and net dividend plots for both the OLS and MM methods. The spread of residuals is wide, with large negative residuals on the left and right side of the plots.

51. It is important to note that the outliers affecting the slope in the net dividend plot are also likely to affect the franking credit slope (θ). This is because, within the multiple

[75] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, p. 31. Also see pages 46-47.

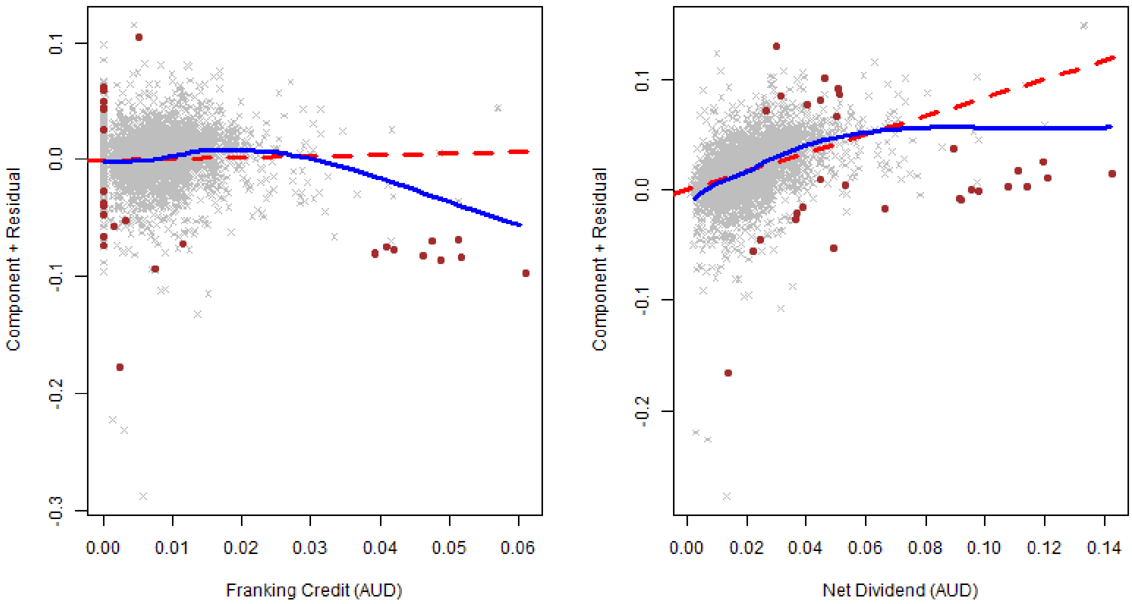
[76] W.S. Cleveland, 'LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression', *The American Statistician*, vol. 35, 1981, p. 54.

regression framework, a change in one parameter estimate can lead to changes in estimates of the other parameters.

52. Figure 12 shows that the most influential points with negative residuals in the OLS plots are spread across high and low net dividend values. The influential points with positive residuals are concentrated above the start of inflection in the nonlinear fit in the net dividend plot as the trend begins to level out.

53. Figure 13 shows that in the MM regression the influential points are located exclusively above the inflection point of the nonlinear regression line in the net dividend plot. This shows that the nonlinearity may have as strong an influence on the robust regression as the OLS regression shown in Figure 12. Comparing the net dividend plot to the franking credit plot in Figure 13 shows that the main influence of nonlinearity is upon the net dividend estimate. This is evident in the majority of high influence points in the franking credit plot being near or on the zero value of the franking credit axis. Points near the zero value will have relatively less influence on a parameter estimate than those further from it due to the absence of an intercept term in the model.

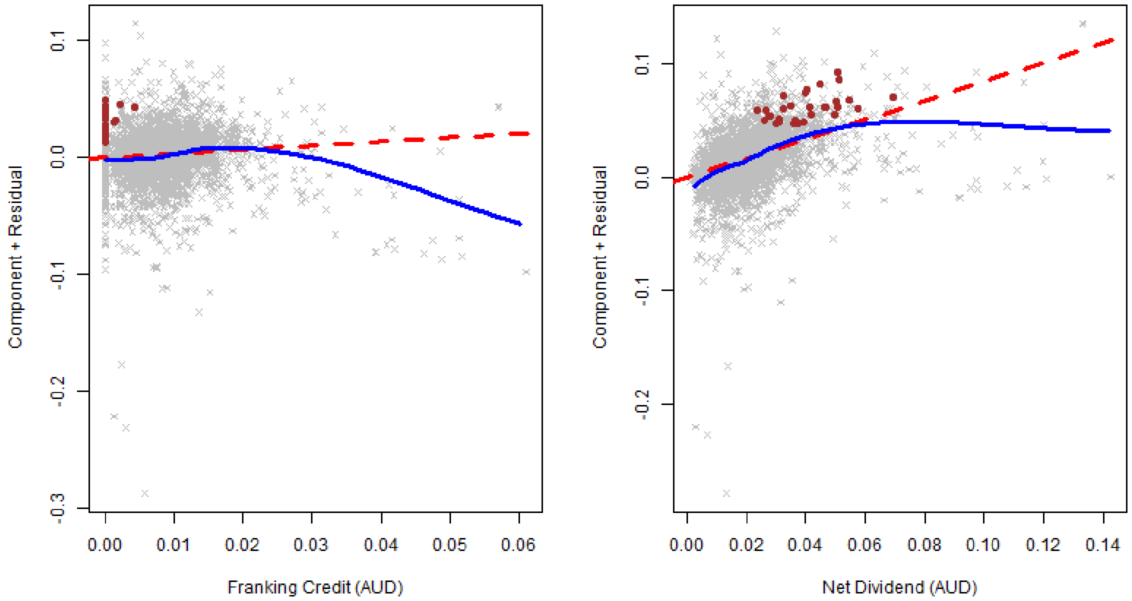
Figure 12 CPR Plots for the OLS estimation of Model 1: market corrected data



Source: ERA Analysis

Note: The linear regression (dashed red line) may be compared with a nonlinear regression (blue line). The 30 most influential observations as identified by the DFBETAs statistic are bolded in each plot, with non-influential observations in grey.

Figure 13 CPR Plots for the MM estimation of Model 1: market corrected data



Source: ERA Analysis

Note: The linear regression (dashed red line) may be compared with a nonlinear regression (blue line). The 30 most influential observations as identified by the DFBETAs statistic are bolded in each plot, with non-influential observations in grey.

54. The key conclusion to be drawn from the CPR plots is that there are different domains of 'behaviour' (or data patterns) that both the OLS and MM regression models are not capturing. The data behaviour that stems from the different domains appear to influence the net dividend coefficient more strongly. Any inaccurate estimate of the net dividends coefficient will then give rise to an inaccurate estimate of theta.

55. The analysis above shows that the simplest division of the different domains of behaviour for model 1 may be defined by the inflection points in the nonlinear regressions for net dividends, at a value of 0.06, and franking credits, at a value of 0.02. Given this behaviour, neither the OLS nor MM regression models can claim they are 'best' fitting.

56. Although disguised somewhat by the scale of the residuals, the slope of the nonlinear fit is greater on average than that of the linear fit for net dividend values less than 0.06, and franking credit values of less than 0.02. This feature is common to all plots in Figure 12 and Figure 13.

A Comment on the SFG Bootstrap Sensitivity Analysis

57. SFG conduct a bootstrap sensitivity analysis that is reproduced for a number of subsequent reports by SFG and Frontier Economics.[77] The bootstrap procedure 'thins' the data by randomly removing five per cent of the full sample a total of $B = 999$ times, before the models are re-estimated:

> To further test the stability of the SFG (2013) theta estimates, we conduct a randomised bootstrapping analysis. To do this, we randomly eliminate five per cent of the sample and re-estimate each of the models using the remaining data. We then repeat this procedure (on the original full sample) another 999 times, yielding 1,000 estimates of theta – each computed after a different 5% of the sample has been removed. This analysis is designed to show how sensitive the estimate of theta might be to removal of 5% of the sample observations.[78]

58. The analysis is designed to show how sensitive the estimates of θ are to the removal of 5% of the sample observations. However, there is little sense as to why one would undertake such an analysis. If 5% of the sample observations are randomly removed then the sample has been thinned, with the thinned sample expected to return a similar estimate of θ. This is because both the full and thinned sample of dividends are drawn from the same sample of data. Moreover, the standard errors for the thinned sample will be expected to be slightly larger than what is estimated for the full sample, unless assumptions underlying the fitted model are not satisfied by the data. For these reasons, it is unusual to undertake this analysis in the way it has been applied by SFG.

59. For example, applying the SFG bootstrap procedure to Vo et al.'s market corrected data results in similar estimates of θ between the Model 1 OLS estimate ($\hat{\theta} = 0.124$) reported in Vo et al. and its bootstrapped mean $\hat{\theta} = 0.119$ reported in Table 14.[79] Moreover, the standard error of $\hat{\theta}$ for the model 1 estimate on market corrected data in Vo et al. is slightly inflated, from 0.11 in the original study to 0.12 in Table 14, due to the reduction in sample size resulting from the random thinning.[80] The 90th percentile confidence band based on the standard error estimate in Table 14, under normality, would span a range of $2 \times z_{0.95} \times SE\left(\hat{\theta}\right) = 2 \times 1.645 \times 0.122 = 0.401$. In comparison, the 90th percentile of the bootstrapped $\hat{\theta}$ spans a range given by 95th percentile less the 5th percentile in Table 14 is 0.403. That is ($0.32 - \left(-0.083\right) = 0.403$. These two range spans are almost equivalent in value.

---

[77] SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21 May 2014, p. 98.

[78] Ibid.

[79] Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013, Table 5.

[80] Ibid, Table 5.

Table 14 Statistics describing the bootstrapped distribution of θ

Statistic	Estimate of θ
Mean	0.119
Median	0.119
Minimum	-0.234
Maximum	0.514
5th Percentile	-0.083
95th Percentile	0.320
Standard Error	0.122

Source: ERA Analysis

60. In contrast, SFG provide a 90 per cent confidence interval that spans a range of $0.208-0.067=0.141$.[81] This range is much less than what the standard errors of $\hat{\theta}$ reported by SFG would suggest, namely $2\times1.645\times0.1946=0.640$.[82] This suggests that SFG have erroneously applied their bootstrap procedure. This is because even if the empirical distribution of the DDO, as represented by the bootstrap procedure, diverges from the normal distribution assumed by OLS regression, the corresponding sampling distributions should largely be similar when the sample size is large (under the Central Limit Theorem). Again, random thinning of a sample should result in slightly larger standard errors, and hence a wider confidence band than the full sample, and not a drastically narrower confidence band as reported by SFG.[83]

61. The inaccuracy of the confidence bands reported by SFG invalidates their assertion that estimates of θ are stable and robust when estimated through dividend drop-off studies. In fact, their result demonstrates the opposite – the standard errors of $\hat{\theta}$ highlights the high uncertainty associated with estimates of θ. Moreover, the question must be asked as to what is actually demonstrated by a sensitivity analysis that randomly removes 5% of the sample. There is no reason to subsample the dividend events given the large sample size of the dividend event data available. In addition, when model misspecification is present (the model fits the data poorly due to breaches of modelling assumptions) then bootstrapping a poorly fitting model has minimal value in terms of informing an appropriate estimate of θ. On this basis the bootstrap sensitivity analysis proposed initially by SFG in May 2014 and later reproduced in subsequent reports by both SFG and Frontier Economics are wholly redundant.

[81] SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21 May 2014, p. 98, Table 3.

[82] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, p. 23.

[83] SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21 May 2014, p. 98, Table 3.

Multicollinearity

62. Franking credits are calculated from the net dividends, and depend on a franking proportion which differs between dividend events (the tax rate correction is proportional).[84] The correlation between the net dividend and franking credit is therefore high (80.8%; without a scaling factor), with a significant proportion of the dividend events having no franking credit assigned, or are fully franked. However, a high correlation does not necessarily imply a level of multicollinearity that impedes parameter estimation.

63. The variance inflation factor (VIF) is a measure of how much multicollinearity inflates the standard error of a regression parameter estimate (as a multiplier of the standard error):

$$VIF\left(\hat{\beta}_p\right) = 1 - R_p^2$$

Where:

$VIF\left(\hat{\beta}_p\right)$ is the variance inflation factor attributable to the p^{th} parameter;

R_p^2 is the R^2 of the regression of the p^{th} predictor variable on the other predictor variables in the equation. Hence, in this study R_p^2 for franking credits may be calculated when franking credits are regressed on net dividends.

64. In general, a predictor variable is said to be highly collinear with other variables in the regression if the square root of its associated VIF is greater than 10.[85] This cut-off constitutes a rule-of-thumb, with other authors recommending a VIF of greater than 5 before multicollinearity among variables is declared. Moreover, the VIF factor may be greater than 10 before the omission of variables from the regression is warranted (even if through a partial down-weighting of the parameter estimates, for example using elastic net procedures that include both ridge and lasso regression).[86]

65. VIF estimates are the same, regardless of whether the response DDO data are market corrected or not. This is because the VIF estimate for θ is based solely on the regression of franking credits on net dividends, and hence is independent of the response. For similar reasons the VIF estimates are consistent between OLS and MM methods, unless the VIF estimate itself is derived through robust (or other) means.

66. When the Model 1 regression was extended, by including an intercept term, the VIF for $\hat{\theta}$ was 2.9. This result implies that the standard error of $\hat{\theta}$ was 1.7 times as large as what it would be if the franking credit was uncorrelated with the net dividend (Table 15). In contrast, as the regression models exclude an intercept term (DDO is strictly

---

[84] Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013, part (iii) of Section (II) Theoretical Background.

[85] M.H. Kutner, C.J. Nachtsheim and J. Neter, *Applied Linear Regression Models*, 4th edn., Chicago, McGraw-Hill/ Irwin, 2004.

[86] R. M. O'Brien, 'A Caution Regarding Rules of Thumb for Variance Inflation Factors', *Quality & Quantity*, vol. 41, 2007, p. 673.

a weighted sum of only the net dividends and franking credits), then the standard error of $\hat{\theta}$ was 8.0 times as large as what it would be if the franking credit was uncorrelated with the net dividend (Table 15). However, VIF are known to be unstable when the intercept term is excluded from a regression.[87]

67. We therefore conclude that the franking credits are moderately to highly collinear with net dividends. The resulting variance inflation may be seen in the large standard errors associated with $\hat{\theta}$ relative to the range of allowable values of 0 to 1, in Vo et al.[88] It can be seen that multicollinearity is slightly stronger for Model 4 than for Models 1 and 3.

68. A high VIF implies that a greater sample size is required to reduce the standard errors of the regression estimates. There is little that can be done to mitigate collinearity, other than acquire a larger sample size through future releases of dividend events. Other options, such as removing either net dividends or franking credits from the regression, or aggregating the net dividends and franking credits into gross dividends, are not feasible solutions as an estimate of θ needs to be directly estimated from franking credit values.

Table 15 Variance inflation factors for the various scaling factors.

Model	VIF($\hat{\theta}$)
1	8.01
2	-
3	8.62
4	10.47

Source: ERA Analysis

Note: No VIF value is returned for model 2 as the scaling factor means that franking credits is the sole predictor variable in the regression model, hence no collinearity with other variables exists.

Heteroskedasticity

69. Heteroskedasticity in a regression setting exists when the residual variance (σ_i^2) of a process depends on the value of the predictor variables through some function $h(\bullet)$ of the predictor variables given by the vector $\underset{\sim}{x}_i$. Residual variance (σ_i^2) is described in this study by the net dividend and the franking credit in the vector $\underset{\sim}{x}_i$:

$$\sigma_i^2 = h\left(\underset{\sim}{x}_i\right)$$

70. Beggs and Skeels model heteroskedasticity as a function of current market capitalisation, gross dividend value, and cum-dividend price at the time of dividend

[87] J. Fox and S. Weisberg, *An R Companion to Applied Regression*, 2nd edn., Thousand Oaks, California, Sage, 2011.

[88] Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013, Table 5.

event i.[89] In contrast, SFG Consulting have modelled heteroskedasticity by applying different scaling factors to the dividend and price data. Moreover, SFG Consulting trial the White (1980) correction for heteroskedasticity of the standard errors, and achieve similar standard error values for a clustering approach based on the characteristics of individual assets.[90] In general, heteroskedasticity does not lead to biased parameter estimates if other assumptions of the linear model are satisfied.

71. Misspecifying the relationship $h(\bullet)$, which is constant valued in the homoscedastic case, leads to inefficiency in the estimation of model parameters, including θ. This invariably leads to an underestimation of the parameter standard errors. Moreover, heteroskedasticity arising from model misspecification may seemingly produce outliers when a misspecified model is fitted to the data.

72. Neither SFG nor Vo et al. report regression diagnostics for whether the prevalent heteroskedasticity is resolved through the application of scaling factors. Nor do these studies trial statistical models that may account for the heteroskedasticity explicitly within the model, and hence arrive at improved standard error estimates.[91]

73. The studentised Breusch-Pagan test is one of a number of possible tests for heteroskedasticity.[92] In applying the test to the OLS estimated models we find that the models scaled by the cum-dividend price have a higher level of heteroskedasticity than those scaled by the net dividend, as indicated by the Breusch-Pagan (BP) statistic (Table 16). Furthermore, the scaling factors scaled by market volatility $\hat{\sigma}_i$ (Models 3 and 4) produce model fits containing a greater level of heteroskedasticity than those models for which the market volatility has not been applied (Models 1 and 2). This means that Model 4 has a statistically significant level of heteroskedasticity as shown by the p-value less than 0.05 in Table 16.

74. Table 16 shows that the market correction appears to deflate the level of heteroskedasticity present in the data. However, this is not to an extent that a different inference about the level of heteroskedasticity present in the data could be drawn.

75. Figure 14 shows regression diagnostics, in the form of scale-location plots, have been applied to Models 1 – 4 for the market corrected data for the 1 July 2001 to 31 July 2012 dividend event data that were previously studied by Vo et al.[93] Consistent with

---

[89] D.J. Beggs and C.L. Skeels, 'Market Arbitrage of Cash Dividends and Franking Credits', *The Economic Record*, vol. 82, 2006, pp.239–252.

[90] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, pp. 21-22.

H. White, 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity', *Econometrica,* vol. 48, 1980, pp. 817-38.

[91] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011.

Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013.

For examples of such diagnostics see: R.D. Cook and S. Weisberg, 'Diagnostics for Heteroskedasticity in Regression', *Biometrika*, vol. 70, 1983, pp. 1-10.

[92] T.S. Breusch and A.R. Pagan, 'A Simple Test for Heteroskedasticity and Random Coefficient Variation', *Econometrica*, vol. 47, 1979, pp. 1287-1294.

[93] A scale-location plot is a plot of the square root of standardized residuals on fitted values from the model. If no heteroskedasticity is present then the mean trend in the scale-location plot should be horizontal.

Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013.

the results from Table 16, models scaled by the cum-dividend price (Models 1 and 4) are more heteroskedastic than those scaled by the net dividend (Models 2 and 3), with Model 4 showing the most extreme heteroskedasticity. Scale-location plots for the robust regression models (MM and LAD) reveal a similar pattern between scaling factor components and heteroskedasticity as for the OLS models (not presented here).

76. It may therefore be assumed that standard errors for the models with cum-dividend prices as a component of the scaling factor S are underestimated relative to the models with net dividend as a component of S.

Table 16 Breusch-Pagan (BP) tests for heteroskedasticity in OLS models.

	Model	BP Statistic	P-value
No market correction	1	3.48	0.06
	2	0.23	0.63
	3	2.71	0.10
	4	82.77	0.00
Market correction	1	3.09	0.08
	2	0.01	0.94
	3	1.06	0.30
	4	60.27	0.00

Source: ERA Analysis

Skewness and Kurtosis

77. Skewness and, to a lesser extent kurtosis, can have an influence on the DFBETAs analysis. It can also introduce bias into regression estimates. Typically, skewness is diagnosed by Q-Q plots, but is also measurable.

78. Any skewness coefficient greater than 1 in absolute value is considered an indication of a highly skewed distribution. High values of kurtosis (greater than three) indicate that the distribution of residuals is leptokurtic. That is, compared to a normal distribution its tails are longer and fatter (having greater probability), with the central peak higher and narrower. This leads to greater variance as the result of infrequent extreme deviations.

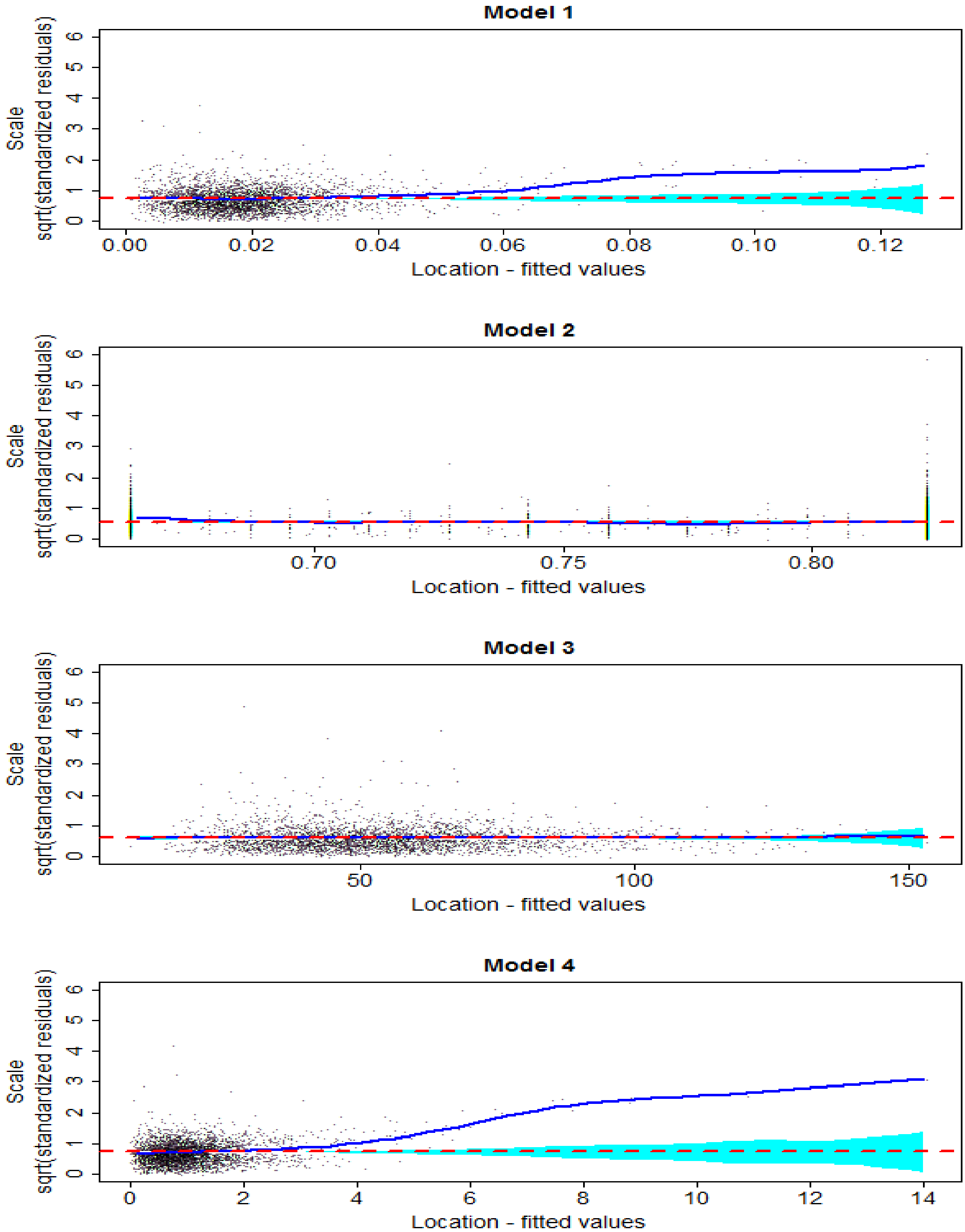
79. Table 17 provides measures of skewness and kurtosis for the OLS estimated models. The key patterns in the data are that the models incorporating net dividend in the scaling factor S are subject to significantly more negative skewness (and greater kurtosis) than those incorporating the cum-dividend price. Moreover, the market correction inflates the skewness slightly. However, in all models the skewness and kurtosis measures indicate that the OLS residuals are both highly skewed and leptokurtic (Table 17).

Table 17 Skewness and kurtosis of residuals from the different OLS models.

	Model	Skewness	Kurtosis
No market correction	1	-1.6	14.7
	2	-11.9	316.3
	3	-5.9	107.0
	4	-2.2	24.1
Market correction	1	-2.5	24.9
	2	-14.9	463.9
	3	-7.7	159.9
	4	-3.5	44.2

Source: ERA Analysis

Figure 14 Scale-location plots for the market corrected Models 1-4



Source: ERA Analysis

Note: Data points are dividend events. The red line indicates a homoskedastic scale value, whereas the dark blue line indicates a heteroskedastic error rate. For homoskedasticity to be assumed then the dark blue line needs to exist fully within the shaded light blue region.

Differences in Estimates of θ

80. Both the Vo et al. and SFG have thus far reported on at least 24 models in total. These models have applied four different scaling factors, three estimation methods (OLS, MM and LAD) and used data with and without a market correction.[94]

81. In studying the marginal means in response to each model choice it can be seen that the largest influence on OLS estimates of θ is the choice of scaling factor evident in the difference between the first and second rows of Table 18 and Table 20.

82. In particular, including the cum-dividend price in the scaling factor results in a significantly lower OLS estimate of θ than when net dividend is included in the scaling factor S (Models 2 and 3). This suggests that large negative residuals are influencing the regression for the cum-dividend price scaling.

Table 18 Mean *θ* as a function of scaling factor components and estimation method

Component of S	OLS	MM	LAD	Robust methods	All methods
cum-dividend price (Model 1 and 4)	0.01	0.33	0.30	0.31	0.21
net dividend (Model 2 and 3)	0.49	0.39	0.46	0.42	0.45
All scaling factors	0.25	0.36	0.38	0.37	0.33

Source: ERA Analysis

Table 19 Mean *θ* as a function of estimation method and market correction

Estimation method	Non-market corrected	Market corrected	All corrections
OLS	0.28	0.22	0.25
MM	0.37	0.34	0.36
LAD	0.43	0.33	0.38
Robust methods	0.40	0.34	0.37
All methods	0.36	0.30	0.33

Source: ERA Analysis

94 Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013, Table 8.

SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, p. 47, Table 10.

SFG have considered three other robust methods (least trimmed squares, M estimation and S Estimation).

The Authority has considered varying the tuning parameter that controls the balance between S and M estimation within the MM method.

Table 20 Mean θ as a function of scaling factor components and market correction

Component of S	Non-market corrected	Market corrected	All corrections
cum-dividend price (Model 1 and 4)	0.23	0.19	0.21
net dividend (Model 2 and 3)	0.50	0.40	0.45
All scaling factors	0.36	0.30	0.33

Source: ERA Analysis

83. The magnitude of difference between the OLS estimates of θ between different scaling models is on average close to 50 per cent (either side of the average in the last column of Table 18). The average OLS estimate from models that scale by the cum-dividend price (0.01) is considerably lower than all other averages in Table 18. This may lead one to ask whether estimates based on this model are infeasibly low, or whether the θ estimate for models with net dividend included in S are too high.

84. The negative skewness for the OLS estimated model 2 and 3 (which include net dividend in S) is greater than for model 1 and 4 (that include cum-dividend price in S). On this basis one would expect a larger number of negative influential outliers in model 2 and 3. One would also expect that these outliers would render the estimates from Model 2 and 3 less reliable than the estimates from model 1 and 4. However, this is not necessarily the case. The relatively lower negative skewness of models 1 and 4 (Table 17) generates highly influential outliers for two reasons:

- the negative skewness in model 1 and 4 is in combined with a greater level of heteroskedasticty, and so skewness at higher franking credit values will have more influence on the models estimates; whereas
- much of the negative skewness in models 2 and 3 occurs at a zero franking credit value, and so have much less influence due the regression intercept being maintained through zero by design.

Despite being subject to greater skewness and kurtosis models 2 and 3 are potentially less influenced by the resulting extreme values than models 1 and 4.

85. The estimates of θ were somewhat sensitive to the market correction (Table 19 and Table 20). On average, the estimate of θ was 0.06 lower for when data were market corrected than for when the data were not market corrected. This makes sense as market returns are on average positive (otherwise there would be no market), and so the DDO values would be expected be lower than without a market correction. Estimates of θ would therefore be reduced with a market correction, on average. This decrease in the θ estimate was consistent across the different estimation methods (Table 19).

86. Estimates of θ were largely consistent for the different models when estimated by the MM method (Table 18), more so than for the other estimation methods. While the MM method was mostly insensitive to the market correction (with a difference of 0.03; Table 19), it was slightly sensitive to the components of scaling factor (with a difference of 0.06; Table 18). The sensitivity of the LAD estimator was between that of the MM and OLS estimators to both the market correction and scaling factor.

Differences in the Standard Error of θ

87. The standard error of θ is greater for OLS regression than for MM methods (Table 21). For OLS estimates the standard error scales with the estimate of θ. This can be seen by comparing the mean estimates in Table 18 to Table 19 where the standard error is greater for models 2 and 3 than for models 1 and 4. The market correction influences the standard error minimally (Table 22) and (Table 23).

88. The standard errors provided for θ remain high. Note that even for MM regression, with a standard error of 0.10 across all models, this equates to a confidence bound for θ that is approximately ± 0.20 about a mean value of 0.36 (Table 18). Much of this imprecision in the standard error estimates may be attributed to multicollinearity between net dividends and franking credits, as well as in the large spread of data values arising from skewness and leptokurticity.

Table 21 Mean standard error of *θ* as a function of scaling factor and estimation method

Component of S	OLS	MM	LAD	Robust methods	All methods
cum-dividend price (Model 1 and 4)	0.11	0.09	0.14	0.11	0.11
net dividend (Model 2 and 3)	0.24	0.11	0.15	0.13	0.17
All scaling factors	0.18	0.10	0.14	0.12	0.14

Source: ERA Analysis

Table 22 Mean standard error of *θ* as a function of estimation method and market correction

Estimation method	Non-market corrected	Market corrected	All corrections
OLS	0.18	0.18	0.18
MM	0.10	0.09	0.10
LAD	0.15	0.13	0.14
Robust methods	0.13	0.11	0.12
All methods	0.15	0.13	0.14

Source: ERA Analysis

Table 23 Mean standard error of *θ* as a function of scaling factor components and market correction

Component of S	Non-market corrected	Market corrected	All corrections
cum-dividend price (Model 1 and 4)	0.12	0.10	0.11
net dividend (Model 2 and 3)	0.17	0.16	0.17
All scaling factors	0.15	0.13	0.14

Source: ERA Analysis

Performance of the existing models

89. Vo et al. reported that:

> The results of the DFBETAs analysis confirm that the estimate of theta is highly sensitive to the choice of the underlying sample of dividend events. Removing just 30 observations from a sample of 3309 can result in a dramatically different estimate of theta. In the course of this process, the value of theta can vary between 0.3 to 0.55. [95]

90. In reply, Frontier have stated that:

> Vo et al (2013) implement a stability analysis known as the DFBETAs approach. This approach differs from the SFG stability analysis in two primary ways:
>
> a) Influential observations are removed one-at-a-time, rather than in pairs; and
>
> b) The stability analysis is only applied in relation to the non-standard approach whereby prices are not corrected for market movements over the ex-dividend day. [96]

91. The market correction has little influence on the sensitivity observed within the DFBETAs analysis. The market correction is important because it results in a marginally lower estimate of θ than if the market correction was omitted.

92. Paired removal within the stability analysis, for reasons explained above (paragraphs 34-45), results in a false impression of smoothness whenever the data are skewed. In contrast, one-at-a-time removal better indicates the distribution of extreme values and their influence on the regression. That said, the SFG stability analysis is both unwarranted because similar results can be generated from the standard DFBETAs that identifies the 30 most extreme outliers. It is also unnecessary because it diagnoses little with regard to violations of the common assumptions of the multiple linear regression model. For these reasons, the SFG stability analysis should not be conducted as part of future dividend drop-off studies.

93. The SFG stability analysis does not reveal or resolve the main data problems in Australian DDO studies. The high variability of the OLS estimates of θ to different scaling factors is an indication of more intrinsic problems. In contrast, the removal of outliers is simply treating a symptom of the intrinsic data problems, rather than designing the analysis to address the non-standard features in the data.

94. A thorough application of regression diagnostics has been presented here, with the diagnostics summarised in Table 12. Critically, each of the four scaling factors have data issues which violate the assumptions of the standard linear model. Overall, when the effects are tallied, Models 1 and 4 are impacted by more of these data issues than Models 2 and 3. This is borne out in the exceptionally low OLS estimates of θ for Models 1 and 4 on average.

---

[95] Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013, part (iii) of Section IV) Analysis.

[96] SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21 May 2014, p. 65, paragraph 462.

Table 24 Summary of violations of the standard linear model for the DDO model given different scaling factors

Model	Outliers (OLS DFBETAs)	Nonlinearity (CPR Plot)[a]	Multicollinearity (VIF)	Heteroskedasticity (Breusch-Pagan)	Skewness	Kurtosis
1	**X**	**X**	*x*	**X**	*x*	*x*
2	**X**				**X**	**X**
3	**X**		*x*		**X**	**X**
4	**X**	**X**	**X**	**X**	*x*	*x*

Source: ERA Analysis

Note: *Nonlinearity was also tested for in comparisons with the broken-stick model detailed in Appendix 3.*

x indicates that the model is affected by the issue. ***X*** *indicates a greater impact than x.*

95. It therefore does not make sense to pick one model over another. For instance, SFG Consulting claim that Model 4 is to be preferred over the other models, following their stability analysis.[97] However, the combined heteroskedasticity and skewness of the Model 4 residuals likely means this model is the least robust of all four models from a diagnostic perspective.

96. The volatility estimate, estimated from excess returns in the year prior to a dividend event, should be selected carefully. For scaling factors with net dividend as a component the inclusion of the standard deviation in excess returns, for market corrected data, has in this instance increased the level of negative skewness while reducing the leptokurtic shape of the distribution. For scaling factors including the cum-dividend price the opposite is true (Table 24). In neither scenario does the inclusion of the standard deviation in excess returns measure resolve either the severe skewness or heteroskedasticity of the data. Including the standard deviation in excess returns demonstrates no clear superiority to scaling factors excluding them. There is no strong reason to include the standard deviation in excess returns within the scaling factor as it is currently constructed.

97. Of interest is the reliability of the robust estimate of θ across the different scaling factors, in particular the MM estimator.[98] If the Model 1 and Model 4 results are discarded then the θ estimate would be 0.37 (Table 18). Moreover, the MM estimator for Models 2 and 3 is associated with lower standard errors (see Table 18) than the other estimators considered.

98. However, the evidence from the regression diagnostics indicate there are potential data issues for which robust regression is not a sufficient solution. All of the DDO studies conducted thus far assume that there is little or no model misspecification

[97] SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21 March 2011, pp. 31.

SFG Consulting, *An Appropriate Regulatory Estimate of Gamma, Report for Jemena gas Networks, ActewAGL, APA, Networks NSW (Ausgrid, Endeavour Energy and Essential Energy), ENERGEX, Ergon, Transend, TransGrid and SA Power Networks,* 21 May 2014, p. 96, paragraph 467.

[98] Reliability in this case is defined here as the extent to which an estimate varies minimally across different model specifications.

error (in the linear regression formula).[99] While robust regression down-weighs extreme values, it is an appropriate model only when the data are linear. The evidence from the CPR plots (Figure 12 and Figure 13) indicates nonlinearity, if not multi-modality of the data (clusters of differently behaving data).

99. Many data points have already been excluded from the data set through the filters that have been applied to the data prior to analysis. It is therefore unclear what further criteria could be applied to filter out potentially confounding data that has not already been removed during the initial filtering process. Outlier removal is therefore not a preferred option, unless unbiased reasons can be supplied as to why the data needs to be removed.

100. Regression models for skewed, leptokurtic data which take heteroskedasticity into account are feasible. However, they assume that the pattern of skewness is in some sense constant. Moreover, the need to exclude an intercept term from the analysis when combined with negatively skewed data requires an appropriate transform of the data, thereby complicating the analysis. Hence, models with an explicit distributional form that accounts for the data features are not a first choice in broadening the types of regression models being considered.

101. The alternative is to apply multi-modal models that assign different regression estimates to different clusters or segments of the data. Of these the broken-stick model and a finite-mixture-of-regressions model is trialled in Appendix 3. In addition, an interaction model including a term representing the interaction between net dividends and franking credits is considered.

[99] N.J. Hathaway and R.R. Officer, *The Value of Imputation Tax Credits,* Working Paper, Melbourne Business School, 2004.

D.J. Beggs and C.L. Skeels, 'Market Arbitrage of Cash Dividends and Franking Credits', *The Economic Record*, vol.82, 2005, pp. 239-252.

SFG Consulting, *Dividend Drop-Off Estimate of Theta, Final Report, Re: Application by Energex Limited (No 2) [2010] ACompT7,* 21st March 2011.

Appendix 3 Alternative Models

1. There is a strong reason to believe that the DDO response to dividends and franking credits is multi-modal, or at least nonlinear (for example, Figure 13). The response to multi-modality can be one of excluding data that correspond to extreme clusters of data (or modes; through robust analysis or outlier removal), or one of modelling the different clusters explicitly. In this manner, the estimates of θ may be said to be dependent on the value of one or more of the predictor variables. Moreover, including an interaction term between franking credits and net dividends constitutes another means of incorporating a varying θ coefficient within a regression model.

2. Hence, we consider three classes of linear regression models that model a varying θ coefficient explicitly:

   - A broken-stick (or segmented) regression, which segments the data along the number line of the predictor variable. The advantage of this approach is that nonlinearities can be modelled as a series of linear segments.
   - A finite-mixture-of-regression models (**FMR**) that segments the data in the combined (mathematical) space defined by all predictor variables. The advantage of this approach is that the identification of data clusters occurs in the joint space, and so is more appropriate for modelling multiple clusters of data rather than smooth nonlinearities.
   - An interaction model, which is simply a multivariate model that includes a linear interaction term representing the interaction between franking credits and net dividends.

Broken-Stick Model

Definition

3. A broken-stick model is one which is piecewise linear, where two or more straight lines are connected at a sequence of break points. More formally, the model for a two segment regression may be defined as:[100]

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 \left(x_i - \psi\right)_+ + \varepsilon_i$$

Where:

$\left(x_i - \psi\right)_+ = \left(x_i - \psi\right) I\left(x_i > \psi\right)$ defines the distance of an observation x_i from a break point ψ whenever $x_i > \psi$;

$I\left(\bullet\right)$ is an indicator function (returns 1 if statement between brackets is true and 0 if false);

β_0, β_1 and β_2 are regression parameters; and

$\varepsilon_i \sim N\left(0, \sigma^2\right)$ are normally distributed residuals with variance σ^2.

4. This model may readily be extended to more than two segments, as well as to the generalized linear model. Estimation requires an initial guess of the unknown breakpoints, with breakpoints and parameter estimates updated in an iterative fashion. The breakpoints are estimated by introducing a 'gap' parameter of the difference between two adjoining straight lines given a breakpoint value. When the algorithm converges then standard errors for the breakpoint estimates are derived using the Delta method.[101]

5. The model estimates can be sensitive to choice of initial breakpoints. However, this simply requires a grid search of initial breakpoint values to arrive at an optimal model that provides sensible parameter estimates. Parameter estimates are highly sensitive to the choice of the number of segments to include in the model. Again, applying a model selection criterion, for example, stepwise selection with Akaike's Information Criterion, results in a model defining a parsimonious set of segments that improves model fit.

Results and Discussion

6. The data for Model 2 was for the most part linear, with the fitted broken-stick model not statistically different from the linear OLS regression (p-value = 0.26). Model 3 data produced a statistically significant model (p-value < 0.001), and while the estimate for θ was feasible for a significant proportion of franking credit values (

[100] V.M.R. Muggeo, (2003) 'Estimating Regression Models with Unknown Break-Points', *Statistics in Medicine*, vol. 22, 2003, pp. 3055–3071.

D. Hinkley. 'Inference in Two-Phase Regression', *Journal of American Statistical Association*, pages 736–743, 1971.

[101] V.M.R. Muggeo, (2003) 'Estimating Regression Models with Unknown Break-Points', *Statistics in Medicine*, vol. 22, 2003, pp. 3055–3071.

$\hat{\theta} = 0.490$ for market corrected data), the estimate of δ was not ($\hat{\delta} = 1.27$). This reflects greater uncertainty in the location of break points when the data are approximately linear, with a consequent reduction in the stability of parameter estimates to changes in the initially selected break points.

7. Data for Models 1 and 4 were clearly nonlinear (Figure 15 and Figure 16). Being able to segment the net dividend variable, and thus allow for nonlinearity, results in $\hat{\theta} = 0.596$ for Model 1 and $\hat{\theta} = 0.767$ for Model 4 when applying the market corrected data (Table 27). Differences between the broken-stick and standard linear models are statistically significant (ANOVA test, p-value <0 .001), with the broken-stick model providing a better model fit.

8. The second segment of each model defines the estimate of θ, and represents approximately 76% to 88% of the data (Table 25 and Table 26).

9. The standard errors for $\hat{\theta}$ are equivalent to that for the MM robust regression (0.11 – 0.14, given $\hat{\theta}$ is larger). The regression diagnostics are improved over the MM regression, insofar as the nonlinearity in the data are taken into account - skewness and kurtosis remain little changed, but the Breusch-Pagan statistic is much reduced for all models (p-value >> 0.05). Only the non-market corrected Model 1 retained a high value for the Breusch-Pagan statistic (p-value = 0.002). This suggests that for Model 4 the severe heteroskedasticity present in the linear regression models is likely a consequence of the nonlinearity in the data.

Figure 15 CPR Plots for the broken-stick model applied to the market corrected Model 1

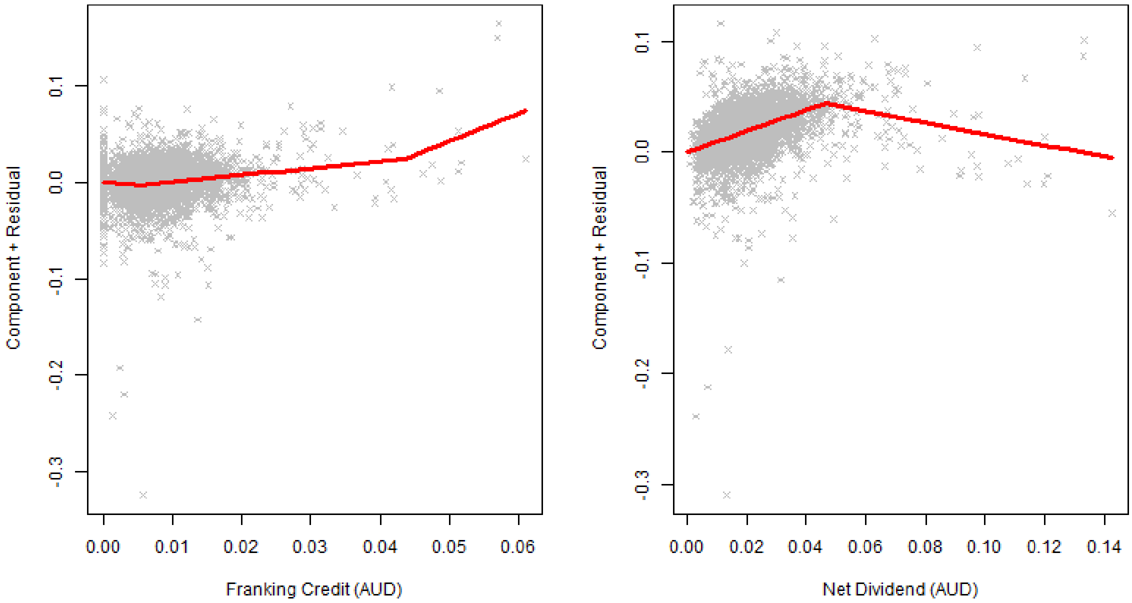
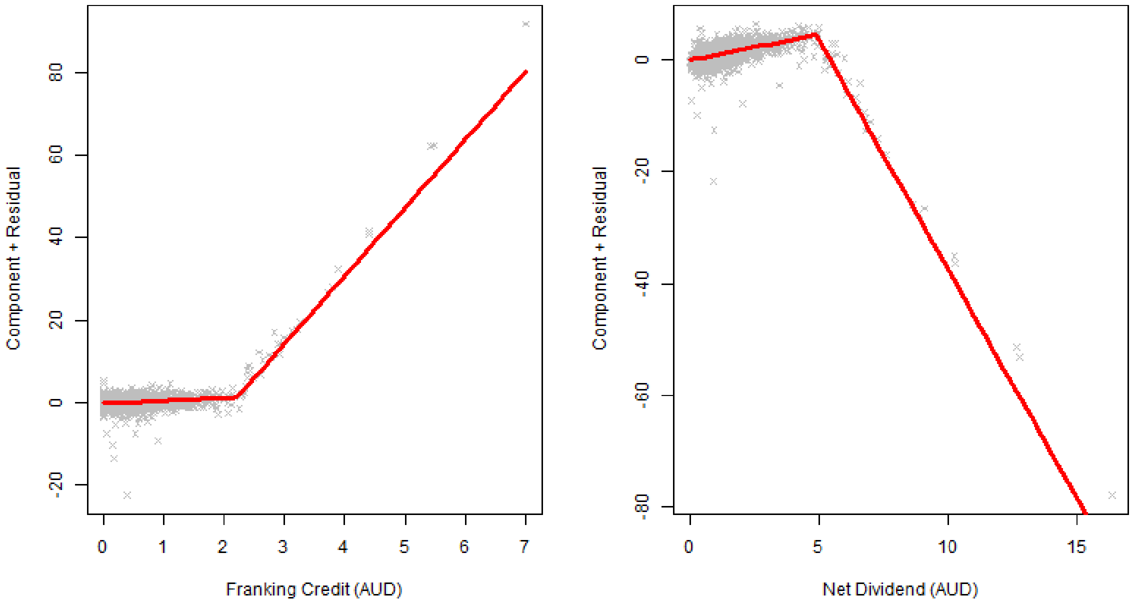


Figure 16 CPR Plots for the broken-stick model applied to the market corrected Model 4



Source: ERA Analysis

10. Estimates of δ from the broken-stick model are greater than for the linear regressions reported by Vo et al. (0.91-0.96; Table 25). Standard errors were equivalent to that of the robust MM regression (0.038-0.049), and the corresponding segment of the regression equated to 97.4% to 98.3% of the data sample (Table 25 and Table 26).

Table 25 Broken-stick estimates of *θ* for non-market corrected data

Model	Parameter	Quantity	Franking Credit Segment 1	Franking Credit Segment 2	Franking Credit Segment 3	Net Dividend Segment 4	Net Dividend Segment 5
1	Slope[a]	Estimate	-2.701	**0.600**[b]	2.591	**0.960**[c]	-0.493
		SE	1.929	0.139	2.654	0.049	0.164
	Range[d]	Estimate	0,0.002	0.002, 0.047	0.047, 0.061	0,0.046	0.046, 0.142
		SE	0,0.001	0.001, 0.003	0.003,0	0,0.002	0.002,0
	Proportion of Sample[e]		0.231	0.768	0.002	0.979	0.021
4	Slope[a]	Estimate	-1.462	**0.719**[b]	1.001	**0.910**[c]	-1.055
		SE	0.749	0.121	0.994	0.043	0.142
	Range[d]	Estimate	0,0.153	0.153, 3.953	3.953, 6.998	0,4.080	4.080, 16.330
		SE	0,0.054	0.054, 0.448	0.448,0	0,0.123	0.123,0
	Proportion of Sample[e]		0.236	0.762	0.002	0.981	0.019

Source: ERA Analysis

Notes: :[a] *provides an estimate of* θ *(segment 2) or* δ *(segment 4), depending on the segment selected.*

[b] *This is the estimate of* θ *derived from segment 2 along the franking credit axis of the broken-stick model. This segment represents the greatest proportion of the sample.*

[c] *This is the estimate of* δ *derived from segment 1 along the net dividend axis of the broken-stick model. This segment represents the greatest proportion of the sample.*

[d] *This is the range gives the values of predictor variable (franking credit or net dividend) at which the break points in the broken-stick model occur. For example, a range estimate of "0,0.002" for segment 1 indicates that the first segment covers the scaled franking credit values from 0 to 0.002, with the standard error in the estimate of the break point being 0.001 (for the observed end points defining the full range of franking credit values the standard error is assumed in this case to be 0).*

[e] *This is the proportion of the sample falling within a segment on a given axis.*

11. The fact that the broken-stick model accounts for nonlinearity in the data, when it is present, better than the standard linear model suggests that a further filtering of the data may potentially be applied to satisfy the assumptions of the linear model. These filters would target for removal exceptionally low values or high values for both franking credits and net dividends. These regions of data are where highly influential observations predominate as a result of the skewness and heteroskedasticity. For example, the market corrected Model 4 suggests scaled franking credits of less than 0.15 and greater than 3.95 exhibit different behaviour to the rest of the sample, along with scaled net dividends of greater than 4.08 (equates to the range estimates in Table 25 and Table 26).

Table 26 Broken-stick estimates of θ for market corrected data

Model	Parameter	Quantity	Franking Credit			Net Dividend	
			Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
1	Slope[a]	Estimate	-1.724	**0.596**[b]	4.879	**0.920**[c]	-0.944
		SE	1.329	0.124	2.050	0.043	0.251
	Range[d]	Estimate	0,0.002	0.002, 0.046	0.046, 0.061	0,0.057	0.057, 0.142
		SE	0,0.001	0.001, 0.002	0.002,0	0,0.003	0.003,0
	Proportion of Sample[e]		0.114	0.881	0.005	0.974	0.026
4	Slope[a]	Estimate	-1.270	**0.767**[b]	1.057	**0.916**[c]	-1.150
		SE	0.614	0.110	0.835	0.038	0.164
	Range[d]	Estimate	0,0.165	0.165, 3.715	3.715, 6.998	0,4.207	4.207, 16.330
		SE	0, 0.051	0.051, 0.358	0.358,0	0,0.120	0.120,0
	Proportion of Sample[e]		0.182	0.808	0.001	0.983	0.011

Source: ERA Analysis

Notes: *:[a] provides an estimate of θ (segment 2) or δ (segment 4), depending on the segment selected.*

[b] This is the estimate of θ derived from segment 2 along the franking credit axis of the broken-stick model. This segment represents the greatest proportion of the sample.

[c] This is the estimate of δ derived from segment 1 along the net dividend axis of the broken-stick model. This segment represents the greatest proportion of the sample.

[d] This is the range gives the values of predictor variable (franking credit or net dividend) at which the break points in the broken-stick model occur. For example, a range estimate of "0,0.002" for segment 1 indicates that the first segment covers the scaled franking credit values from 0 to 0.002, with the standard error in the estimate of the break point being 0.001 (for the observed end points defining the full range of franking credit values the standard error is assumed in this case to be 0).

[e] This is the proportion of the sample falling within a segment on a given axis.

12. If the range estimates from the broken-stick model are used to filter out further data so as to maintain the linearity assumption then complications arise in terms of specifying the intercept. For the current data, the broken-stick results in a lower intercept for the second segment assigned to the franking credits than would occur if either the intercept is excluded or included. This is because the first segment has negative slope, and has zero intercept. Furthermore, the broken-stick model does not apply to all scaling of the data (does not possess generalisability), and the model is sensitive (does not possess reliability) to the specification of the number and value of the initial break points in the search for optimally valued break points. Consequently, the broken-stick model should not be applied in a regulatory context. However, the model does clearly demonstrate that:

- The scaling by cum-dividend price should not be considered in final estimates of θ, as the broken-model corrects for the heteroskedasticity arising from the nonlinearity of the data. As the data do not satisfy key assumptions of the linear model then a linear model should not be applied.

- A higher estimate of θ may be achieved if a better fitting model is specified. A better fitting model will account for one or more of the violations of standard assumptions of the linear model.

Finite Mixture-of-Regressions

Definition

13. A finite mixture-of-regressions model fits a number of regression models simultaneously to the data.[102] Instead of assuming that the data are nonlinear, as with the broken-stick model, the model assumes that there are multiple 'latent' clusters present in the data, each of which are linear. The *latent variable* here is a variable that assigns the data to the different classes, but which is not known *a priori* to model estimation. Part of the model estimation process is to assign data to different latent classes in a way that maximises model fit. Hence, a mixture-of-regressions can be applied to data where observations originate from various groups and the group affiliations are not known, or to provide approximations for multi-modal processes.

14. The basic framework for mixture-of-regression models firstly defines a density distribution g of the predictor variables:[103]

$$g_{\eta}\left(D_i, FC_i\right) = \sum_{m=1}^{M} \omega_m \phi\left(DDO_i \mid \delta_m D_i + \theta_m FC_i; \sigma_m^2\right)$$

Where:

ϕ is the standard normal distribution, defined by a parameter vector of parameters η;

η is the parameter vector $\eta = \left(\delta_1, \ldots, \delta_M, \theta_1, \ldots, \theta_M, \sigma_1^2, \ldots, \sigma_M^2\right)$;

ω are positive weights summing to unity; and

m indexes one of M latent distributions (in this case normally distributed, bivariate regression models, defined by the regression parameters $\left(\delta_m, \theta_m, \sigma_m^2\right)$.

15. Mixture-of-regression models are typically estimated through some variant of the EM algorithm (under a frequentist approach), or are estimated through Bayesian

---

[102] R.E. Quandt, 'A New Approach to Estimating Switching Regressions', *Journal of the American Statistical Association*, vol. 67, 1972, pp 306-310.

B.S. Everitt and D.J. Hand, *Finite Mixture Distributions*, London, Chapman and Hall, 1981.

D. Titterington, A. Smith and U. Makov, *Statistical Analysis of Finite Mixture Distributions*. Hoboken, New Jersey, John Wiley and Sons, 1985.

McLachlan G, Peel D, *Finite Mixture Models*, John Wiley and Sons Inc, 2000.

[103] B. Gruen and F. Leisch. 'FlexMix Version 2: Finite Mixtures with Concomitant Variables and Varying and Constant Parameters', *Journal of Statistical Software*, vol. 28, 2008, pp. 1-35

T. Benaglia, D. Chauveau, D.R. Hunter and D. Young, "mixtools: An R Package for Analyzing Finite Mixture Models." *Journal of Statistical Software*, vol. 32, 2009, pp. 1–29.

methods. Standard errors of the model parameters may be estimated through a bootstrap procedure for EM estimation.[104]

Results and Discussion

16. Here, we estimate a two cluster regression for each model, which was the optimal number of clusters M for Models 2, 3, and 4. For Model 1 a five cluster model may feasibly be estimated, but is not considered here for comparison purposes (the five cluster model returns similar slope coefficients to the two cluster model for the relevant cluster). The cluster that returns feasible θ and δ estimates has been reported in Table 27.

17. Regressions for each of the two clusters were defined with zero intercept.

18. The estimates of θ are very close to those of the robust MM regression estimates, with standard errors for the mixture-of-regression estimates only slightly larger.[105] The mixture-of-regression demonstrates similar patterns in the parameter estimates as to the robust MM regression estimates, with non-market corrected data having higher θ estimates and lower δ estimates than the market corrected data.

19. The mixture-of-regression model demonstrates generalisability (can be applied to all models) and reliability (reproduces similar estimates under different scenarios on the whole). However, the functionality of the mixture-of-regression software employed for model estimation is not fully flexible. The software is constrained insofar as the zero intercept needs to be enforced for all of the cluster regressions. In practice, one would want to enforce a zero intercept for the main cluster of data, yet allow non-zero intercepts for the other clusters.

20. In effect, constraining the regressions of the non-dominant clusters to have zero intercept effectively filters out the extreme data from the dominant cluster (Figure 17). This is equivalent to a robust MM regression which assigns zero or little weight to the extreme data, and hence it may be of no surprise that parameter estimates between the two classes of regressions are similar. This is achieved by the second cluster having a higher variance estimate than the main cluster. Extreme points either side of the main regression are thus assigned relatively more weight (probability of membership to the cluster) by the regression with the higher variance estimate.

21. The mixture-of-regressions model in this scenario does little to counter nonlinearity in the data, as evidenced in part by the mismatch between the regression lines not fitting well the corresponding data cluster in the CPR plot for Model 4 (Figure 17). Overall, there is little to suggest from the mixture-of-regression models that the data expresses itself as multiple, discrete clusters.

[104] A.P. Dempster, N.M. Laird and D.B. Rubin, 'Maximum Likelihood from Incomplete Data Via the EM-Algorithm', *Journal of the Royal Statistical Society B*, vol. 39, 1977, pp. 1–38.

[105] Compared with values reported in Table 5 of Vo, D., Gellard, B., Mero, S., *'Estimating the Market Value of Franking Credits, Empirical Evidence From Australia' Conference Paper*, Australian Conference of Economists 2013, 2013.

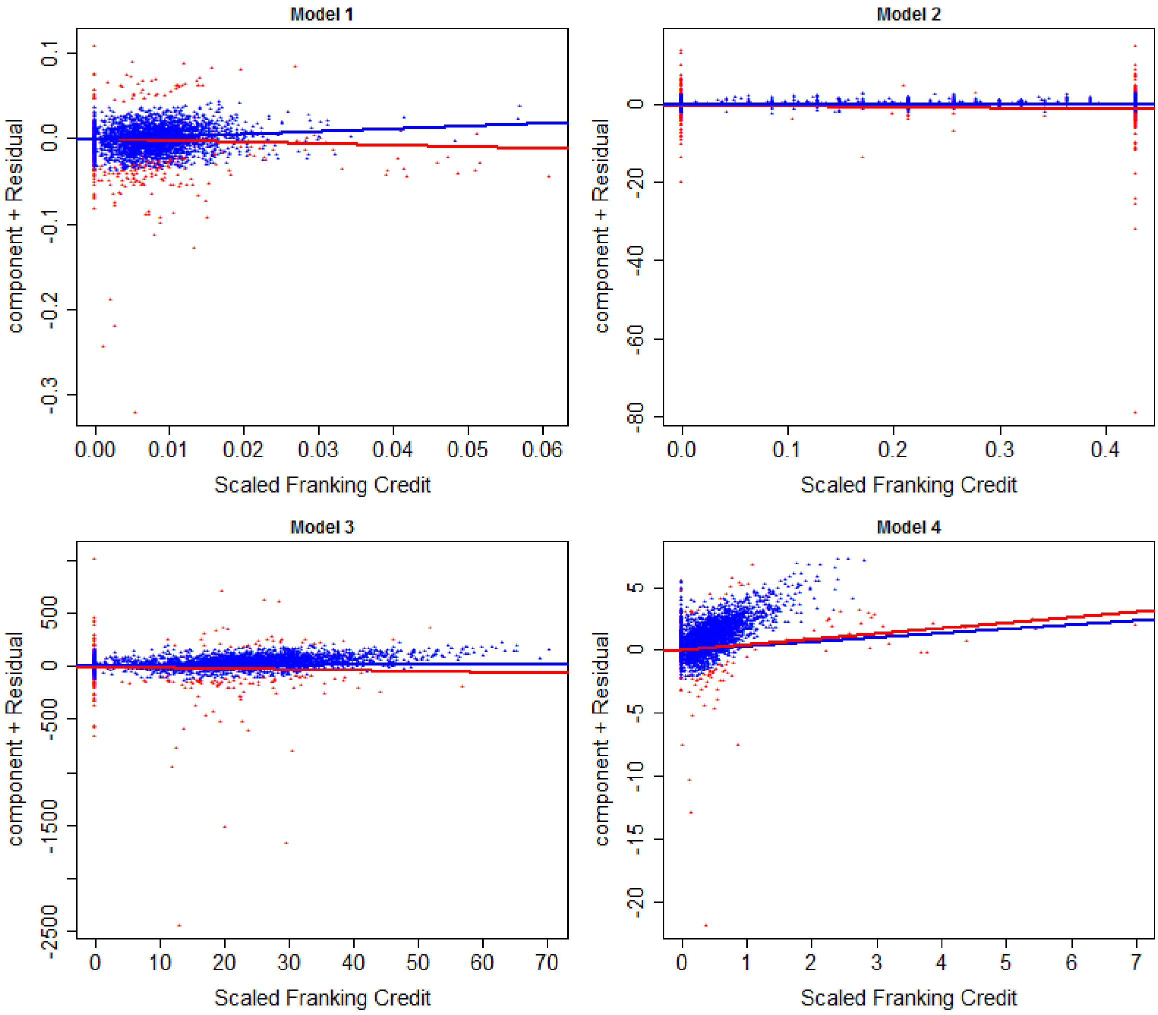
Table 27 Mixture-of-regression results

Correction	Model	Value	Franking Credit θ	Net Dividend δ	Residual Variance σ	Proportion of Data
Non-market corrected	**0**	**Estimate**	0.325	0.077	0.020	0.947
		Std. Error	0.181	0.015		
	1	**Estimate**	0.507	0.820	0.014	0.889
		Std. Error	0.119	0.047		
	2	**Estimate**	0.365	0.796	0.893	0.932
		Std. Error	0.129	0.050		
	3	**Estimate**	0.448	0.791	50.9	0.942
		Std. Error	0.127	0.050		
	4	**Estimate**	0.378	0.869	0.921	0.953
		Std. Error	0.105	0.041		
Market Corrected	**0**	**Estimate**	0.250	0.928	0.020	0.971
		Std. Error	0.115	0.047		
	1	**Estimate**	0.308	0.926	0.014	0.926
		Std. Error	0.095	0.038		
	2	**Estimate**	0.365	0.831	0.855	0.942
		Std. Error	0.120	0.047		
	3	**Estimate**	0.412	0.844	47.7	0.948
		Std. Error	0.117	0.046		
	4	**Estimate**	0.342	0.925	0.868	0.970
		Std. Error	0.082	0.032		
Mean – non-market corrected[a]			0.424	0.819	-	0.929
Mean – market corrected[a]			0.357	0.850	-	0.947
Total Mean[a]			0.391	0.882	-	0.938

Note: This is for the cluster returning feasible estimates of θ and δ (out of two in total)

[a] The mean value excludes Model 0 estimates.

Figure 17 CPR plots for mixture-of-regressions applied to the market corrected data



Note: The blue data represent the main cluster of data from which θ and δ estimates are acquired, and may be differentiated from the red, non-dominant cluster. The coloured lines represent the regressions on the scaled franking credit for the corresponding cluster.

An Interaction Model

Definition

22. In addition to nonlinearity and/or clustering in the data, a further feature that may be specified in the linear regression model is an interaction between net dividends and franking credits in the prediction of DDO values. Delivering a causal explanation as to why such interactions occur if observed in the data is not part of the current scope. The interaction model set out here will be treated simply as a black-box, much like the broken-stick and mixture-of-regressions models, that assists in understanding the features of the data.

23. The interaction model may be defined as:

$$\frac{DDO_i}{S} = \frac{\delta D_i}{S} + \frac{\theta FC_i}{S} + \frac{\rho D_i \times FC_i}{S} + \varepsilon_i$$

Where:

$\hat{\rho}$ is the coefficient of the interaction term between D_i and FC_i , with all other terms defined as previously shown.

Results and Discussion

The estimates of θ scale with the net dividend amount. Hence, with a negative value for $\hat{\rho}$ then the $\hat{\theta}$ reported in is the maximum value which is realised when net dividend values are low. A lower estimate of θ will, in this case, be observed when net dividends are large. This holds true for Models 1 and 4 (at least for MM estimation), whereas the reverse is true for Models 2 and 3 due to a positive $\hat{\rho}$.

24. Estimates of θ in the interaction model are thus a function of the scaled net dividend. This function can be described in terms of quantiles of the scaled net dividend (Table 29 expressed as percentiles). The dependence of $\hat{\theta}$ on the scaled net dividend is a detraction of the model in terms of delivering a single value of $\hat{\theta}$ for the purposes of the Weighted Average Cost of Capital calculation. However, in almost all cases the interaction model provides a statistically significant improved fit of the model to the data over and above the linear model without interaction.

25. The interaction model does nothing to resolve the skewness or kurtosis in the data. However, heteroskedasticity in the data appears to be reduced significantly, given observed Breusch-Pagan statistics (not reported here).

26. The OLS and MM estimation methods produced quite dissimilar estimates for each scaling factor applied to the data. For each estimation method parameter estimates varied widely with respect to which scaling factor was applied.

27. Despite this variation, it may be observed that the interaction model provides a reasonably sensible estimate of θ for the majority of the data. Only when net dividends exceed their 90th percentile (or equivalently the less than the 10th percentile for Model 3; Table 29) are unreasonable estimates of θ observed (as in Model 4; Figure 18).

28. As an example, when parameter estimates are considered across Models 1, 2 and 4 then $\hat{\theta}$ ranges from 0.25 to 0.61, given the 0th to 90th percentiles of the range of net dividend values for $\hat{\theta}$ estimated through MM regression (Table 29). The mid-point of this range is 0.43.

Table 28 Interaction model results

Model	Estimation Method	No Market Correction $\hat{\delta}$	No Market Correction $\hat{\theta}$	No Market Correction $\hat{\rho}$	Market Correction $\hat{\delta}$	Market Correction $\hat{\theta}$	Market Correction $\hat{\rho}$
0	OLS	0.688 (0.010)	0.666 (0.274)	-0.100 (0.016)	0.721 (0.007)	0.815 (0.197)	-0.109 (0.011)
	MM	0.692 (0.001)	0.561 (0.014)	0.208 (0.001)	0.921 (0.001)	0.179 (0.013)	0.398 (0.001)
1	OLS	0.823 (0.045)	0.605 (0.130)	-11.0 (1.32)	0.835 (0.042)	0.624 (0.121)	-11.1 (1.23)
	MM	0.823 (0.045)	0.365 (0.102)	-0.493 (1.04)	0.896 (0.033)	0.275 (0.094)	0.790 (0.952)
2[a]	OLS	0.568 (0.107)	0.528 (0.281)	0.077 (0.201)	0.663 (0.105)	0.351 (0.275)	0.112 (0.197)
	MM	0.792 (0.043)	0.387 (0.113)	-0.018 (0.081)	0.845 (0.041)	0.344 (0.106)	0031 (0.076)
3	OLS	0.642 (0.085)	0.055 (0.305)	0.007 (0.003)	0.744 (0.081)	-0.252 (0.292)	0.009 (0.003)
	MM	0.791 (0.043)	0.087 (0.156)	0.005 (0.001)	0.856 (0.040)	0.047 (0.145)	0.004 (0.001)
4	OLS	0.856 (0.042)	0.668 (0.114)	-0.176 (0.009)	0.867 (0.039)	0.720 (0.105)	-0.180 (0.008)
	MM[b]	0.869 (0.034)	0.215 (0.091)	0.054 (0.007)	0.920 (0.032)	0.616 (0.085)	-0.155 (0.007)
Mean[c]	OLS	0.749	0.600	- [d]	0.788	0.565	- [d]
	MM	0.811	0.620	- [d]	0.863	0.581	- [d]
	All	0.780	0.610	- [d]	0.826	0.573	- [d]

Note: Standard errors are reported in brackets

[a] *Numerical instability was observed in deriving the model 4 estimates using MM regression on the market corrected data. The estimates provided here have a lower AIC than the alternatives to which the MM algorithm converged.*

[b] *This model has been specified as* $\frac{DDO_i}{D_i} = \delta + \frac{\theta FC_i}{D_i} + \rho FC_i$. *This is a plausible specification, but not the only specification by which to capture an interaction between* FC *and* D.

[c] *The mean values exclude Model 0 (with the highest OLS estimate of* θ *for the market corrected data) and Model 3 (with the lowest OLS estimate of* θ *for the market corrected data).*

[d] *A mean value of* $\hat{\rho}$ *is has no interpretation when averaged over models defined by different scaling factors.*

Table 29 Estimates of θ as a function of the quantiles of the scaled net dividend, applied to the market corrected data

Model	Estimation Method	Quantiles of the Net Dividend						
		0	10th	25th	50th	75th	90th	100th
0	OLS	0.814	0.811	0.809	0.803	0.793	0.768	-14.5
	MM	0.180	0.191	0.199	0.220	0.259	0.349	55.9
1	OLS	0.613	0.520	0.471	0.401	0.327	0.230	-0.955
	MM	0.275	0.282	0.286	0.291	0.296	0.303	0.387
2	OLS	0.343	0.345	0.345	0.347	0.350	0.357	4.69
	MM	0.343	0.345	0.345	0.347	0.350	0.357	4.75
3	OLS	-0.171	0.068	0.155	0.268	0.396	0.543	1.251
	MM	0.086	0.201	0.243	0.198	0.359	0.430	0.771
4	OLS	0.715	0.643	0.592	0.518	0.419	0.298	-2.23
	MM	0.611	0.550	0.505	0.442	0.357	0.253	-1.91

Source: ERA Analysis

Conclusions

29. The impetus to fit alternative regression models to the dividend event data was driven by the poor performance of the existing models in terms of regression diagnostics. Importantly, each of the broken-stick, mixture-of-regressions and interaction models improved model fits in a statistically significant manner. Moreover, these alternative models have resulted in a wider range of estimates.

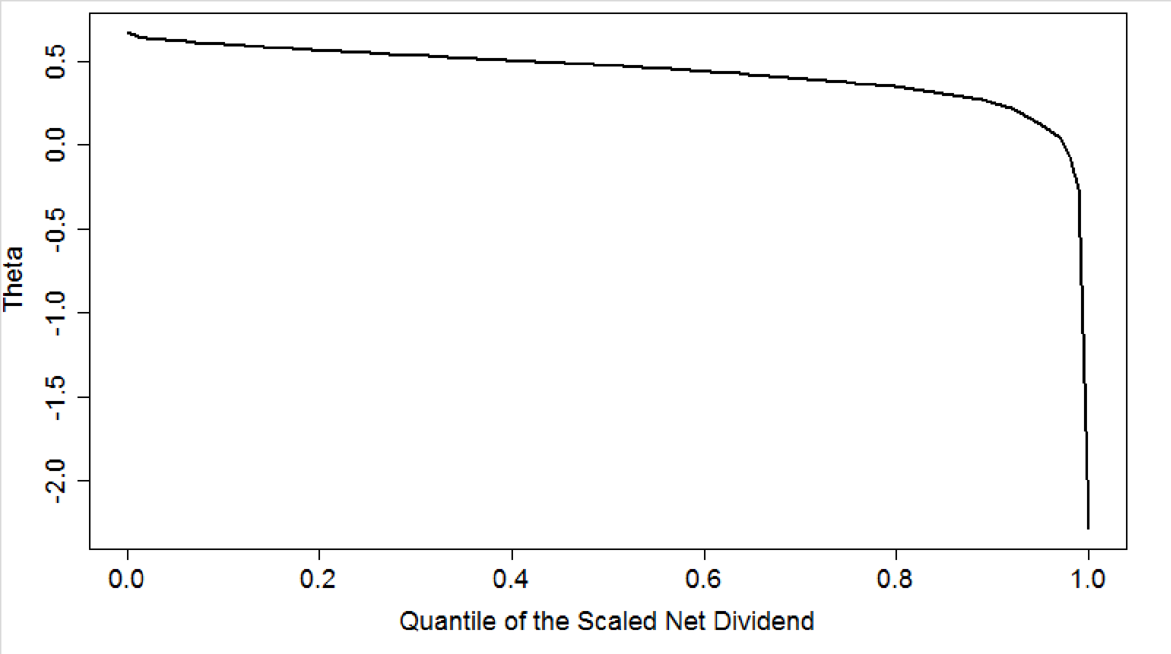
30. The broken-stick model was applicable to models where the cum-dividend price was included in the scaling factor (Models 1 and 4). These models provide an estimate of θ of at least 0.6 for both uncorrected and market corrected data for much of the data range. This was achieved by representing the nonlinearity in the data as piecewise-linear functions that are defined separately for the net dividend and the franking credit.

31. The mixture-of-regression models returned estimates of θ of approximately 0.35 for market corrected data and 0.42 for non-market corrected data (Table 27). These results are similar to results from MM estimation for the standard linear model. It was noted that there may be a need to relax the assumption that all sub-dominant clusters need have a zero intercept, so as to more accurately identify different clusters of extreme data. Otherwise, the mixture-of-regression models were found to be performing in a similar manner to the MM regression in down-weighting extreme observations by assigning those extreme observations to sub-dominant clusters from which θ estimates are effectively ignored.

32. The interaction model estimates averaged at near 0.43 (excluding Model 3) for 90 per cent of the data, but performed inconsistently between models with different scaling factors. Interpretation is more difficult, insofar as the estimate of θ is dependent on the net dividend value. Estimates of θ for high net dividend values (greater than the

90th percentile of the net dividend, in general) were observed to take on impractical values.

Figure 18 Estimates of *θ* for the market corrected model 4 as a function of the quantiles of the scaled net dividend



Source: ERA Analysis

33. In effect, each of these models truncated the data (as does MM regression) to remove extreme data points. For the broken-stick model separate linear regressions were fitted to different segments of the data along the number line of the predictor variables, thereby excluding data from the main segment. For the mixture-of-regressions specification of an alternative cluster effectively removes data from the dominant cluster. For the interaction model we applied a heuristic to remove the most extreme 10 per cent of net dividend values to derive a range of θ estimates.

34. None of these models did much to mitigate the effects of skewness or kurtosis, although the broken-stick and interaction model did well in mitigating heteroskedasticity when it was present (in models 1 and 4).

35. In general, robust MM regression is still asymptotically consistent when skewness is present, and provides more efficient estimators (the variance of the parameters is minimised) than OLS, particularly when high skewness and kurtosis are present.[106] However, when data are highly skewed and where the sample size is sufficiently large (as is the case with the current dividend event data) then applying a parameterised distribution to the skewed data, such as the skewed-t distribution, will likely provide a more consistent and efficient estimator of θ.[107]

36. Moreover, the evidence presented here suggests that an optimal model (one that is not poorly specified given the data) will likely not only assume explicitly a skewed distribution of the data, but include interaction and/or nonlinear term(s) to describe the

[106] T.A. Maravina, *'Tests for Differences between Least Squares and Robust Regression Parameter Estimates and Related Topics'*, Doctoral Dissertation, University of Washington, 2012, p. 166.

[107] Ibid.

data. That is, an appropriately specified model will, due to the complexities of the data, be more complex than the models presented here.

37. Regardless, the data are severely skewed, more so for the market corrected data than for the uncorrected data (Table 17). Consequently, deriving a suitably well specified model that produces consistent estimates across the different scaling factors may not be guaranteed. This means there is no compelling reason to choose one model over another to derive a point estimate for θ. As there is no model that is a clear 'winner' in terms of model performance, then it makes more sense to state a range of θ estimates rather than a point estimate. Picking a single model may lead to a perverse outcome if the model is poorly specified. An example of this is the preference of SFG/Frontier Economics for Model 4, where by standard regression diagnostics this model was perhaps the worst performing of the scaling factors examined here.