Royal Commission
into Aged Care Quality and Safety

THE COST OF RESIDENTIAL AGED CARE

TECHNICAL SUPPLEMENTARY REPORT 2: COST FRONTIER ANALYSIS OF AUSTRALIAN RESIDENTIAL AGED CARE FACILITIES

APPENDICES

RESEARCH PAPER 9

AUGUST 2020

The Royal Commission into Aged Care Quality and Safety was established by Letters Patent on 8 October 2018. Replacement Letters Patent were issued on 6 December 2018, and amended on 13 September 2019 and 25 June 2020.

The Honourable Tony Pagone QC and Ms Lynelle Briggs AO have been appointed as Royal Commissioners. They are required to provide a final report by 26 February 2021.

The Royal Commission releases consultation, research and background papers. This research paper has been prepared by the University of Queensland for the information of Commissioners and the public. The views expressed in this paper are not necessarily the views of the Commissioners.

This paper was published on 27 August 2020 in six parts: the main report; appendices to the main report; Technical Supplementary Report 1; appendices to Technical Supplementary Report 1; Technical Supplementary Report 2; and appendices to Technical Supplementary Report 2 (this part).

ISBN 978-1-921091-32-2 (online)

Technical Supplementary Report 2: Cost frontier analysis of Australian residential aged care facilities

APPENDICES

Contents

Tables

Figures

Appendix A

Data diagnostics: Identifying and analysing outliers in cost and output data

A-1 Introduction

Before proceeding with the data analysis, it is typically considered advisable to check and explore the data quality and, in particular, the existence of outliers and their nature. This is important because the results may be heavily influenced by the quality of data in general and in particular by outliers that may contaminate the sample and potentially distort the results. While there is never 100% immunity from problems in the data, it is better to reduce the likelihood of the problems as much as possible before proceeding with the main analysis. Outlined below is the findings in this respect and the proposed steps.

A-2 Output

Figure A-1 shows the histogram and boxplot of the (unadjusted) total bed days. Based on the boxplot, 89 outliers were identified. Compared to the whole sample, the group of these outliers has a relatively higher proportion of metro facilities and also a relatively higher proportion of facilities which have lower quality ratings as shown in Table A-1 (this uses the 5 level index version which was later collapsed to 3).

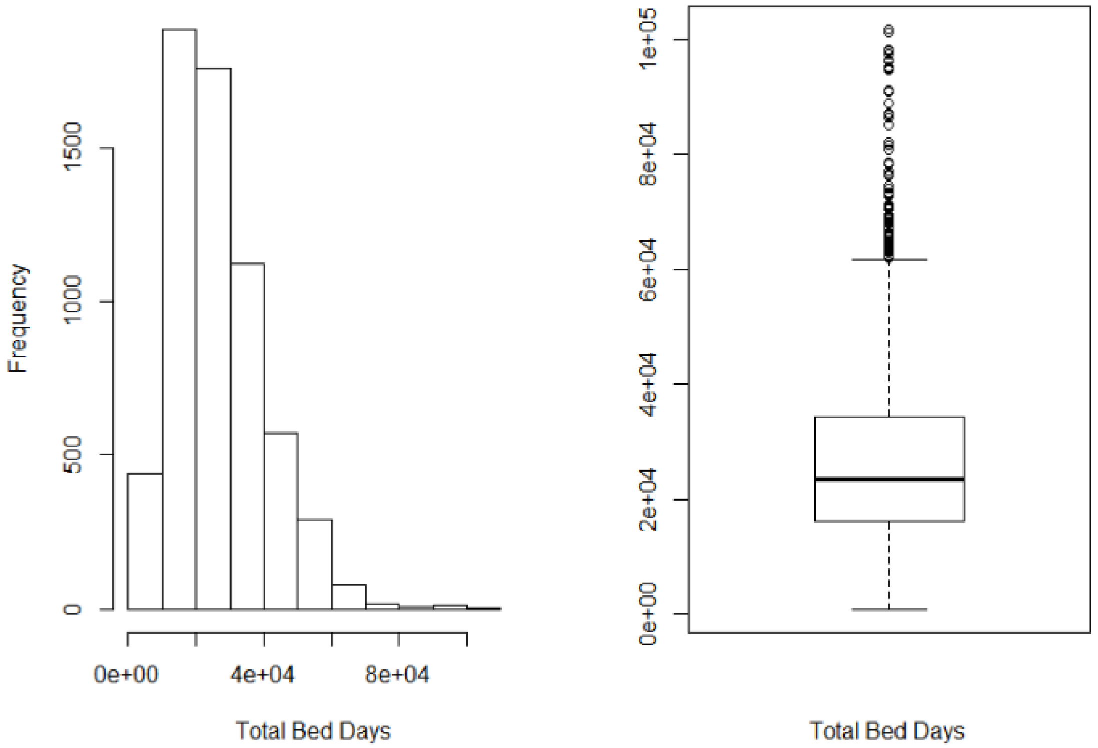

Figure A-1. Histogram and boxplot of the (unadjusted) total bed days

Table A-1. Characteristics of facilities in the group of total bed day outliers

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Total	89	6188	100%	100%	1%	2.08	2.00
Provider type							
Not-for-profit	76	5025	85%	81%	2%	2.07	1.97
For-profit	11	889	12%	14%	1%	2.12	2.16
Government	2	274	2%	4%	1%	2.57	1.95

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Location							
Metropolitan	78	3811	88%	62%	2%	2.06	2.02
Regional	11	2298	12%	37%	0%	2.22	1.97
Remote	0	79	0%	1%	0%	NA	1.77
Size							
Small (<30 beds)	0	367	0%	6%	0%	NA	1.68
30+ beds	89	5821	100%	94%	2%	2.08	2.02
Quality							
Level 1 (highest)	0	658	0%	11%	0%	NA	1.77
Level 2	41	3272	46%	53%	1%	2.06	1.98
Level 3	32	1681	36%	27%	2%	2.07	2.10
Level 4	13	424	15%	7%	3%	2.20	2.09
Level 5 (lowest)	3	153	3%	2%	2%	2.04	2.11

NA: not applicable; No.: number; Obs: observations

Before looking at adjusted output, the total bed days adjusted by RVU individual casemix, we examined the distributions of the RVU individual casemix adjustment. Its histogram and boxplot are shown in Figure A-2. The distribution of casemix is close to bell-shaped, with the mean by construction at 1, and the minimum value at about 0.4, while the highest value at about 1.6. Most of the observations are around the mean.

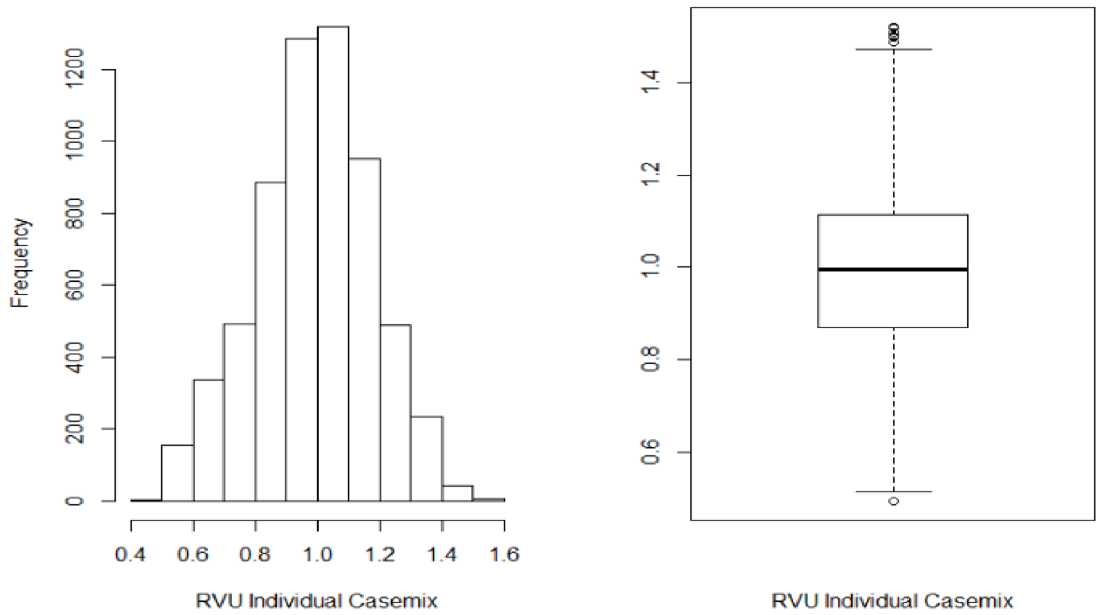

Figure A-2. Histogram and boxplot of relative value unit (RVU) individual casemix

For the adjusted output, based on its boxplot (Figure A-3), 114 outliers were identified. Compared to the whole sample, the group of these outliers has a relatively higher proportion of metro facilities and also higher proportion of facilities which have low quality as shown in Table A-2. At this stage, these outliers were not removed.

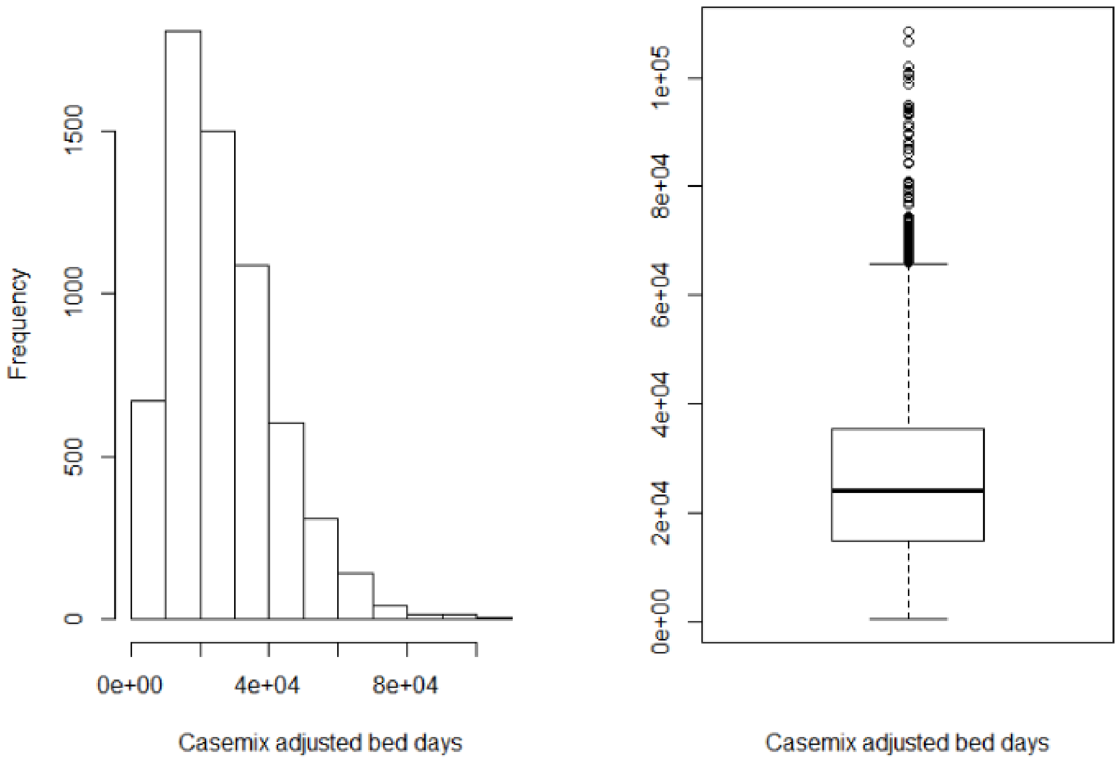

Figure A-3. Histogram and boxplot of casemix adjusted total bed days

Table A-2. Characteristics of facilities in the group of casemix adjusted total bed day outliers

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Total	114	6188	100%	100%	2%	2.26	2.00
Provider type							
Not-for-profit	92	5025	81%	81%	2%	2.25	1.97
For-profit	20	889	18%	14%	2%	2.31	2.16
Government	2	274	2%	4%	1%	2.57	1.95
Location							
Metropolitan	102	3811	89%	62%	3%	2.27	2.02
Regional	12	2298	11%	37%	1%	2.23	1.97
Remote	0	79	0%	1%	0%	NA	1.77
Size							
Small (<30 beds)	0	367	0%	6%	0%	NA	1.68
30+ beds	114	5821	100%	94%	2%	2.26	2.02
Quality							
Level 1 (highest)	0	658	0%	11%	0%	NA	1.77
Level 2	46	3272	40%	53%	1%	2.21	1.98
Level 3	45	1681	39%	27%	3%	2.28	2.10
Level 4	16	424	14%	7%	4%	2.39	2.09
Level 5 (lowest)	7	153	6%	2%	5%	2.27	2.11

NA: not applicable; No.: number; Obs: observations

A-3 Total cost

For the total cost, based on its boxplot (Figure A-4), 142 outliers were identified. Compared to the whole sample, the group of these outliers has a relatively higher proportion of metro facilities and also relatively higher proportion of facilities which have lower quality as shown in Table A-3.

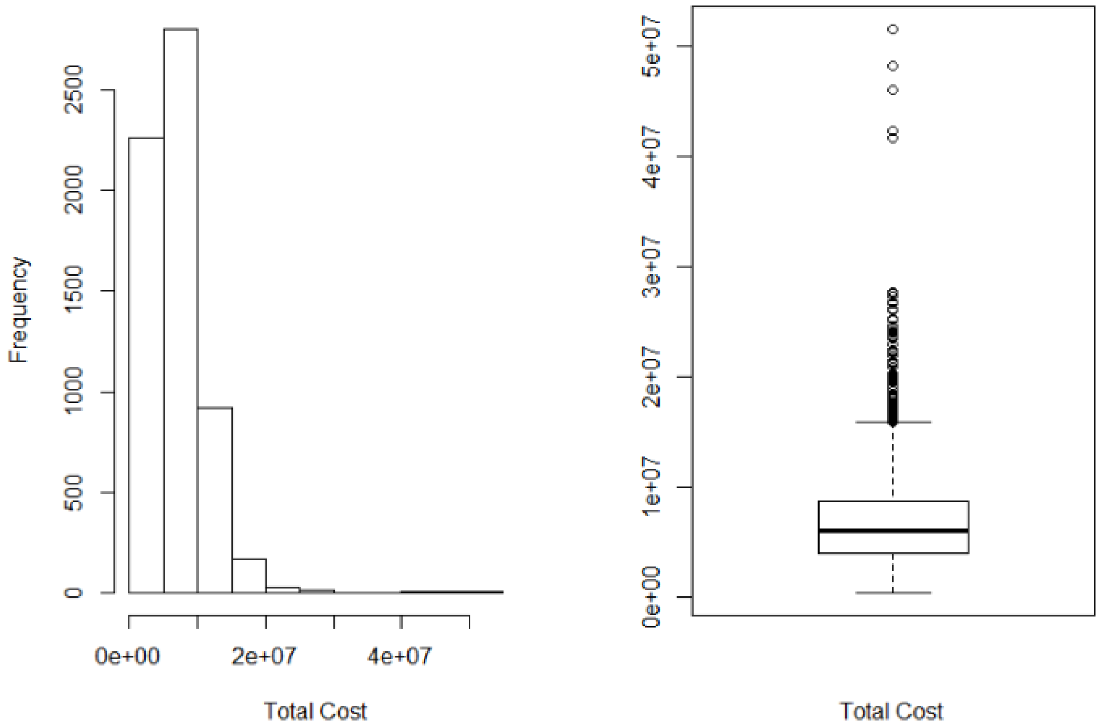

Figure A-4. Histogram and boxplot of total cost

Table A-3. Characteristics of facilities in the group of total cost outliers

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Total	142	6188	100%	100%	2%	2.14	2.00
Provider type							
Not-for-profit	111	5025	78%	81%	2%	2.13	1.97
For-profit	26	889	18%	14%	3%	2.16	2.16
Government	5	274	4%	4%	2%	2.21	1.95
Location							
Metropolitan	124	3811	87%	62%	3%	2.12	2.02
Regional	18	2298	13%	37%	1%	2.23	1.97
Remote	0	79	0%	1%	0%	NA	1.77
Size							
Small (<30 beds)	0	367	0%	6%	0%	NA	1.68
30+ beds	142	5821	100%	94%	2%	2.14	2.02

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Quality							
Level 1 (highest)	0	658	0%	11%	0%	NA	1.77
Level 2	64	3272	45%	53%	2%	2.05	1.98
Level 3	49	1681	35%	27%	3%	2.23	2.10
Level 4	16	424	11%	7%	4%	2.18	2.09
Level 5 (lowest)	13	153	9%	2%	8%	2.14	2.11

NA: not applicable; No.: number; Obs: observations

Of the 142 observations identified as outliers in the boxplot of total cost, 101 observations were identified as outliers in the boxplot of adjusted output (accounting for 71% of outliers in the total cost and 88% of outliers in the adjusted output). This is natural since it requires a high cost to produce high level of outputs. It is worth noting that the group of outliers in both the output and total cost has relatively higher proportions of metropolitan and lower-quality facilities than the average. These outliers have not been removed but may be removed in a sensitivity analysis.

A-4 Average cost

For the average cost, both cost per bed day (unadjusted) and cost per casemix adjusted bed day were analysed. Figure A-5 shows a histogram and boxplot of cost per bed day. Based on the boxplot of cost per bed day, two groups of outliers were identified: extremely low average cost outliers and extremely high average cost outliers. The composition of higher average cost outliers is summarised in Table A-4. From the table, it can be seen that the proportions of regional and remote facilities, governmental facilities, higher-quality facilities, and small facilities in this group of outliers are relatively higher than the average in the whole sample.

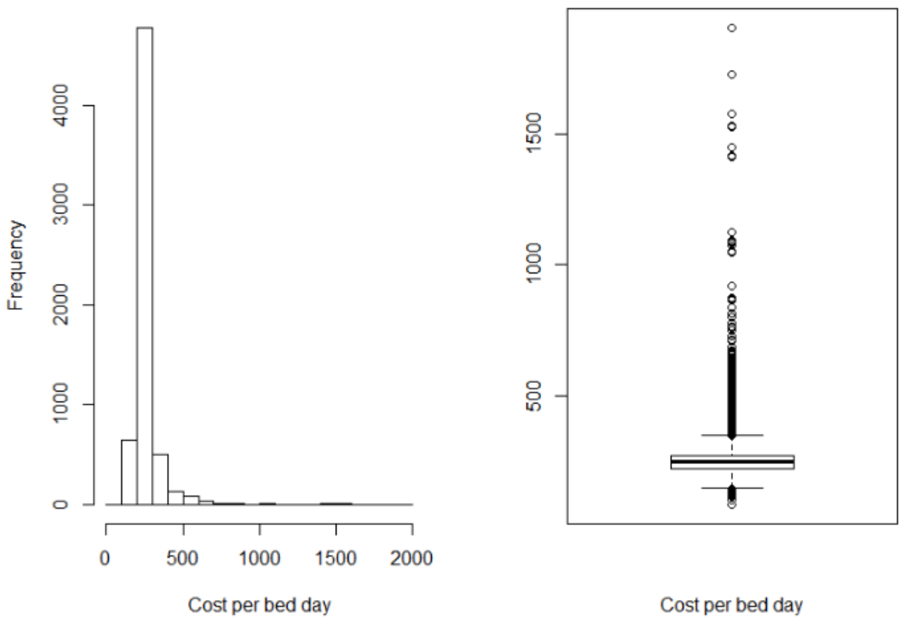

Figure A-5. Histogram and boxplot of cost per bed day

Table A-4. Characteristics of facilities in the group of extremely high cost per bed day outliers

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Total	367	6188	100%	100%	6%	2.09	2.00
Provider type							
Not-for-profit	258	5025	70%	81%	5%	2.07	1.97
For-profit	38	889	10%	14%	4%	2.09	2.16
Government	71	274	19%	4%	26%	2.17	1.95
Location							
Metropolitan	160	3811	44%	62%	4%	2.04	2.02
Regional	184	2298	50%	37%	8%	2.14	1.97
Remote	23	79	6%	1%	29%	2.02	1.77
Size							
Small (<30 beds)	49	367	13%	6%	13%	2.01	1.68
30+ beds	318	5821	87%	94%	5%	2.10	2.02
Quality							
Level 1 (highest)	61	658	17%	11%	9%	2.05	1.77
Level 2	184	3272	50%	53%	6%	2.01	1.98
Level 3	80	1681	22%	27%	5%	2.25	2.10
Level 4	27	424	7%	7%	6%	2.16	2.09
Level 5 (lowest)	15	153	4%	2%	10%	2.18	2.11

No.: number; Obs: observations

The characteristics of low-average cost outliers are shown in Table A-5. Although the average casemix of this group is substantially lower than the average casemix of the sample, with only 32 observations, it is difficult to make any conclusion about the pattern of this group. A more focused case study on each of these observations may be required to analyse and understand their uniqueness. Some of them may represent a 'know-how' on keeping the average cost much lower than others or may reveal some errors in measurement/ reporting that might be useful to realise and cross-check with other observations.

Table A-5. Characteristics of facilities in the group of extremely low cost per bed day outliers

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Total	32	6188	100%	100%	1%	1.41	2.00
Provider type							
Not-for-profit	25	5025	78%	81%	0%	1.42	1.97
For-profit	5	889	16%	14%	1%	1.23	2.16
Government	2	274	6%	4%	1%	1.72	1.95
Location							
Metropolitan	21	3811	66%	62%	1%	1.33	2.02
Regional	11	2298	34%	37%	0%	1.57	1.97
Remote	0	79	0%	1%	0%	NA	1.77

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Size							
Small (<30 beds)	3	367	9%	6%	1%	1.56	1.68
30+ beds	29	5821	91%	94%	0%	1.39	2.02
Quality							
Level 1 (highest)	5	658	16%	11%	1%	1.16	1.77
Level 2	23	3272	72%	53%	1%	1.43	1.98
Level 3	1	1681	3%	27%	0%	2.28	2.10
Level 4	3	424	9%	7%	1%	1.38	2.09
Level 5 (lowest)	0	153	0%	2%	0%	NA	2.11

NA: Not applicable; No.: number; Obs: observations

The histogram and boxplot for the cost per casemix adjusted bed day are shown in Figure A-6. Based on the boxplot of cost per casemix adjusted bed day, again, two distinctive types of outliers are identified: extremely low average cost outliers and extremely high average cost outliers.

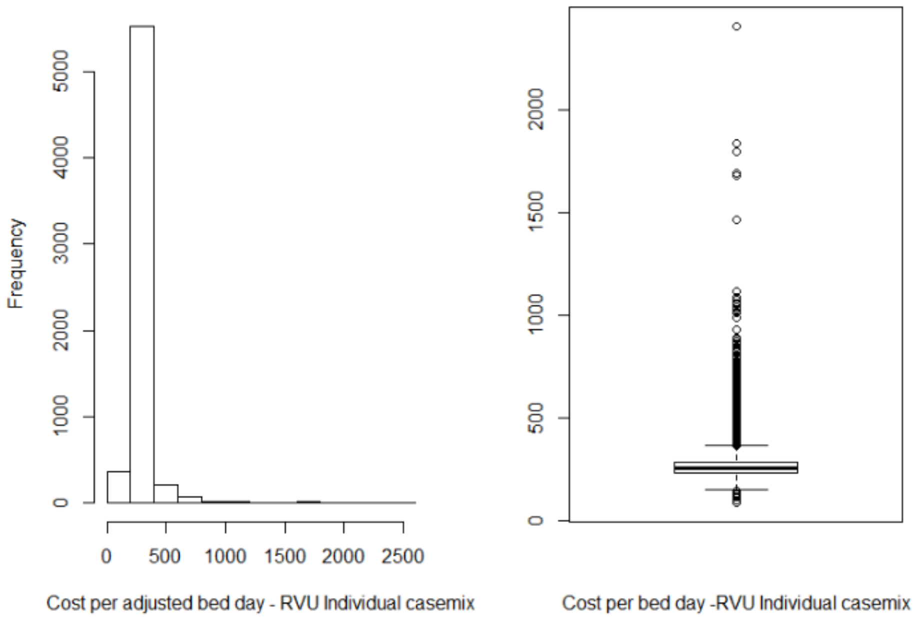

Figure A-6. Histogram and boxplot of cost per casemix adjusted bed day

The composition of higher average cost outliers is provided in Table A-6. From the table, it can be seen that the proportions of regional and remote facilities, government facilities, and small facilities in this group of outliers are relatively higher than the average in the sample. The low average cost outliers group only includes seven observations.[A]

[A] These observations are facilities 1967, 3333, and 4257 in the financial year 2014/2015, 3397, 3546, 3599 in the financial year 2017/2018, and 2906 in the financial year 2018/2019.

Table A-6. Characteristics of facilities in the group of extremely high cost per adjusted bed day outliers

	No. of obs in the group	No. of obs in the sample	Group proportion	Sample proportion	Proportion of group in sample	Average group casemix	Average sample casemix
Total	391	6188	100%	100%	6%	1.84	2.00
Provider type							
Not-for-profit	290	5025	74%	81%	6%	1.82	1.97
For-profit	35	889	9%	14%	4%	1.85	2.16
Government	66	274	17%	4%	24%	1.91	1.95
Location							
Metropolitan	145	3811	37%	62%	4%	1.78	2.02
Regional	212	2298	54%	37%	9%	1.92	1.97
Remote	34	79	9%	1%	43%	1.63	1.77
Size							
Small (<30 beds)	91	367	23%	6%	25%	1.60	1.68
30+ beds	300	5821	77%	94%	5%	1.91	2.02
Quality							
Level 1 (highest)	85	658	22%	11%	13%	1.68	1.77
Level 2	200	3272	51%	53%	6%	1.79	1.98
Level 3	72	1681	18%	27%	4%	2.08	2.10
Level 4	22	424	6%	7%	5%	2.09	2.09
Level 5 (lowest)	12	153	3%	2%	8%	1.89	2.11

No.: number; Obs: observations

In the following analyses, the outliers that are 2.5% in each tail of the distribution of average cost (cost per adjusted bed days) are removed. It results in a trimmed sample of 5,880 observations (308 observations are dropped), which remains a large sample relative to most studies in the literature. For the trimmed sample, the histogram and boxplot of the average cost are shown in Figure A-7.

Even after trimming, the boxplot shows many more outlying observations in the right tail of the distribution. Therefore, additional trimming of that tail might be appropriate and can be considered as a sensitivity analysis. In the characteristics of facilities of the trimmed sample (Table A-7), it can be seen that the trimmed sample represents the original sample well.

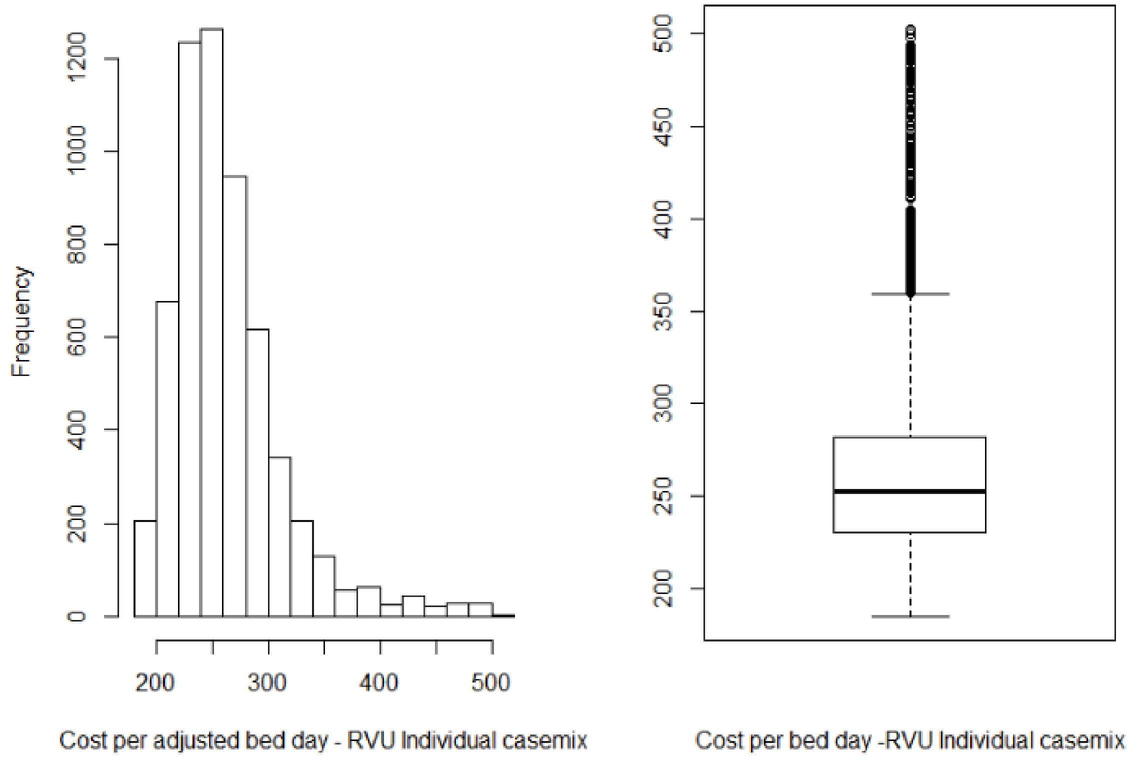

Figure A-7. Histogram and boxplot of cost per casemix adjusted bed day for trimmed sample

Table A-7. Characteristics of facilities in the trimmed sample

	No. of obs in the trimmed sample	No. of obs in the whole sample	Trimmed sample proportion	Whole sample proportion	Proportion of trimmed sample to whole sample	Average trimmed sample casemix	Average whole sample casemix
Total	5880	6188	100%	100%	95%	2.00	2.00
Provider type							
Nor-for-profit	4819	5025	82%	81%	96%	1.97	1.97
For-profit	812	889	14%	14%	91%	2.13	2.16
Government	249	274	4%	4%	91%	1.98	1.95
Location							
Metropolitan	3608	3811	61%	62%	95%	2.01	2.02
Regional	2212	2298	38%	37%	96%	1.97	1.97
Remote	60	79	1%	1%	76%	1.87	1.77
Size							
Small (<30 beds)	318	367	5%	6%	87%	1.68	1.68
30+ beds	5562	5821	95%	94%	96%	2.01	2.02
Quality							
Level 1 (highest)	624	658	11%	11%	95%	1.76	1.77
Level 2	3080	3272	52%	53%	94%	1.97	1.98
Level 3	1637	1681	28%	27%	97%	2.10	2.10
Level 4	396	424	7%	7%	93%	2.07	2.09
Level 5 (lowest)	143	153	2%	2%	93%	2.09	2.11

No.: number; Obs: observations

A-5 Super-efficient outliers diagnostic using data envelopment analysis for the trimmed sample

Data envelopment analysis was used as a tool to diagnose the data to determine super-efficient units, which have a substantially lower cost compared to its peers and examine their influence on the distribution of efficiency scores of the trimmed sample.

We utilise the input-oriented data envelopment analysis models to estimate technical efficiency of facilities. The production technology includes one input and output. The input is the *total cost* and the output is the casemix adjusted bed days (Y*). We employ both *constant returns to scale* and *variable returns to scale* data envelopment analysis estimators.

The *constant returns to scale data envelopment analysis* estimator of input-oriented technical efficiency for a facility with total cost, TC^i, casemix adjusted bed days, Y^{*i}, can be obtained by

$$\widehat{TE}^{i}_{CRS}(TC^i, Y^{*i}) \equiv \min_{\theta, z^1, \ldots, z^n} \theta \qquad (A.1)$$

Subject to

$$\sum_{k=1}^{n} z^k Y^{*k} \geq Y^{*i}, \qquad (A.2)$$

$$\sum_{k=1}^{n} z^k TC^k \leq \theta TC^i, \qquad (A.3)$$

$$\theta \geq 0, z^k \geq 0, k = 1, \ldots n. \qquad (A.4)$$

The *variable returns to scale data envelopment analysis* estimator of input-oriented technical efficiency for a facility with total cost TC^i casemix adjusted bed days Y^{*i} can be obtained by

$$\widehat{TE}^{i}_{VRS}(TC^i, Y^{*i}) \equiv \min_{\theta, z^1, \ldots, z^n} \theta \qquad (A.5)$$

Subject to

$$\sum_{k=1}^{n} z^k Y^{*k} \geq Y^{*i}, \qquad (A.6)$$

$$\sum_{k=1}^{n} z^k TC^k \leq \theta TC^i, \qquad (A.7)$$

$$\sum_{k=1}^{n} z^k = 1, \qquad (A.8)$$

$$\theta \geq 0, z^k \geq 0, \qquad k = 1, \ldots n. \qquad (A.9)$$

For more details and refers on data envelopment analysis, refer to Sickles and Zelenyuk (2019).

First, both data envelopment analysis models (*constant returns to scale* and *variable returns to scale*) were run with the trimmed sample. The distributions of estimated efficiency scores for the trimmed sample using both data envelopment analysis models appear to be much more reasonable, with the majority of facilities being relatively efficient (Figure A-8), as compared to the case of the whole sample (i.e. before trimming the outliers) (Figure A-9).

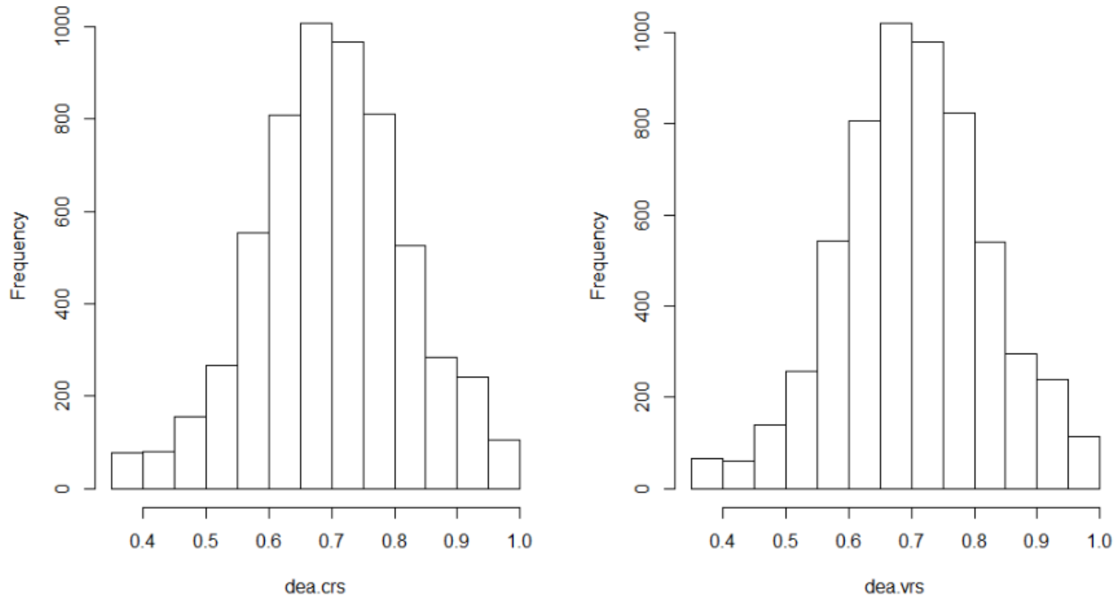

Figure A-8. Histograms of estimated efficiency scores for the trimmed sample using data envelopment analysis for constant returns to scale (DEA-CRS) and variable returns to scale (DEA-VRS)

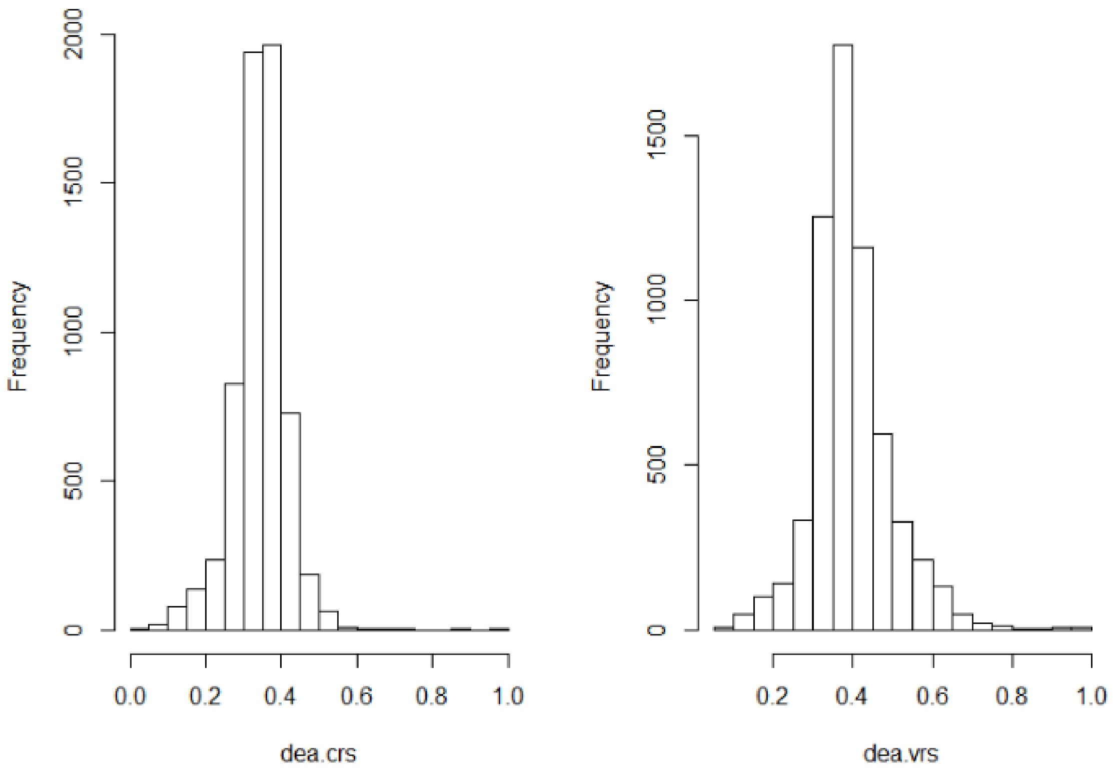

Figure A-9. Histograms of estimated efficiency scores for the whole sample using data envelopment analysis for constant returns to scale (DEA-CRS) and variable returns to scale (DEA-VRS)

For a further investigation (a sensitivity analysis), the data envelopment analysis models (*variable returns to scale)* was run several times, each time the observations lying on the estimated frontiers were identified and excluded from the sample to start the next run. The distributions of efficiency scores after each iteration are in order as illustrated in Figure A-10. The distributions of estimated efficiency scores are stable across iterations, which can confirm that super-efficient outliers are not an issue for the trimmed sample.

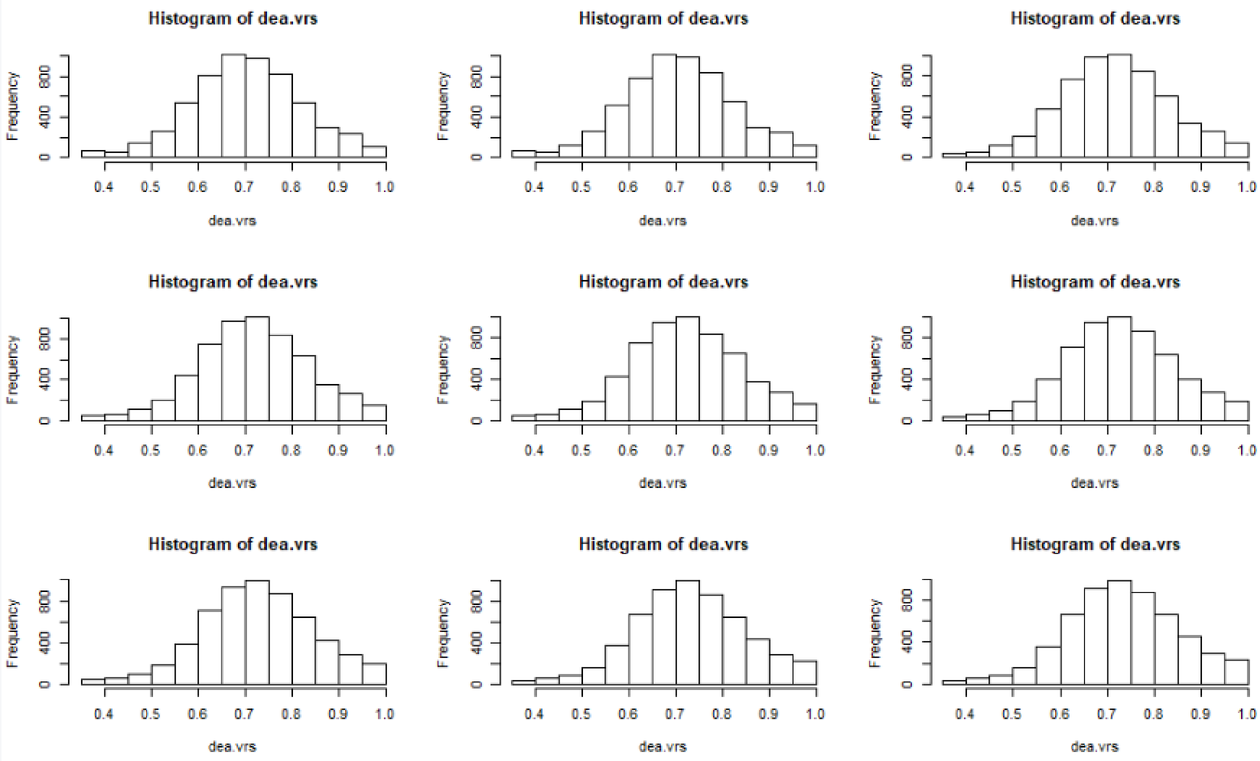

Figure A-10. Histograms of data envelopment analysis for variable returns to scale (DEA-VRS) estimated efficiency scores in the first nine iterations

Appendix B
Robustness check of estimated inefficiency: Semi-parametric least squares stochastic frontier analysis

B-1 Theoretical framework

A limitation of all the stochastic frontier analysis models that have been applied is the specific parametric assumptions imposed on the distributions of error components and the functional relationship for the frontier.

In this section, non-parametric least square methods for stochastic frontier models, developed by Simar, Van Keilegom and Zelenyuk (2017) (hereafter SVKZ), are applied. The goal is to investigate the sensitivity of the estimated inefficiency with respect to the parametric assumptions on the functional form of the frontier and on the distribution of the statistical noise.

Adapting the SVKZ model, the cost function can be formulated as

$$TC_i = m(Y_i, Z_i) + U_i + V_i, \qquad i = 1, \ldots, n, \tag{52}$$

where, TC_i is the total cost, Y_i is the output and Z_i represents other control variables, and $m(Y_i, Z_i)$ is the cost frontier, which is completely unknown to the researchers. V_i is statistical noise, which is assumed to have zero mean (i.e. $E(V_i|Y_i, Z_i) = 0$) and positive finite variance (i.e. $Var(V_i|Y_i, Z_i) \in (0, \infty)$). Meanwhile, U_i is the inefficiency term following some one-sided distribution, with some positive mean (i.e. $(U_i|Y_i, Z_i) = \mu_U$) and positive finite variance (i.e., $Var(U_i|Y_i, Z_i) \in (0, \infty)$). As in other SFA models, U_i and V_i are also assumed to be independent, conditionally on (Y_i, Z_i).

To be able to estimate the inefficiency scores, the parametric assumption on the distribution of inefficiency needs to be made. For example,

$$U_i|Y_i, Z_i \sim \mathcal{N}^+(0, \sigma_U^2(Y_i, Z_i))$$

which can be understood as a local assumption. This is because it is not assumed that any parametric functional form of how U_i depends (conditionally) on Y_i, Z_i via its variance or its mean (after the truncation).

Define

$$\varepsilon_i = V_i + U_i - \mu_U(Y_i, Z_i), \tag{55}$$

where

$$\mu_U(Y_i, Z_i) = E(U_i|Y_i, Z_i),$$

and

$$r_1(Y_i, Z_i) = m(Y_i, Z_i) + \mu_U(Y_i, Z_i). \tag{56}$$

Equation (52) can be written as

$$TC_i = r_1(Y_i, Z_i) + \varepsilon_i. \tag{57}$$

Since $E(\varepsilon_i|Y_i, Z_i) = 0$, standard nonparametric methods (e.g. local polynomial least squares) can be used to estimate $r_1(Y_i, Z_i) = E(TC_i|Y_i, Z_i)$, which can be referred to as the average cost function.

The residuals from the estimation, $\hat{\varepsilon}_i$, can then be utilised to nonparametrically estimate the third moment of the composed error, $r_3(Y_i, Z_i) = E(\varepsilon_i^3|Y_i, Z_i)$. The conditional mean of inefficiency can be obtained by plugging the nonparametric estimate of the third moment into the following relationships

$$\sigma_U^3(Y_i, Z_i) = \sqrt{\frac{\pi}{2}}\left(\frac{\pi}{4-\pi}\right) r_3(Y_i, Z_i), \tag{58}$$

and

$$\mu_U(Y_i, Z_i) = \sqrt{\frac{2}{\pi}}\sigma_U(Y_i, Z_i) \tag{59}$$

The main advantage of the SKVZ model is that it does not impose any parametric assumptions on the cost frontier as well as the distribution of statistical error. Therefore, it provides a robust estimator of the frontier and inefficiency.

Meanwhile, a limitation of this approach in the context of this project is that being similar to the Aigner, Lovell and Schmidt (1977) model (hereafter ALS77), it also does not account for the panel nature of the data, except for the annual effects. However, since the annual effects are accounted via the discrete kernels, they are modelled in a non-parametric manner as well (i.e. not just like shifts of the intercept but also potentially affecting the slopes of all the repressors). This way is more flexible than in parametric fixed effects models, however it does not replace the role of those models.

B-2 Empirical results

Here, the inefficiency of the industry is estimated using the SVKZ model and results are compared with those obtained by using the ALS77 model to investigate the sensitivity of the stochastic frontier analysis models to the parametric assumptions. For ease of comparison, the so-called 'grand cost frontier' (i.e. only output and year dummies are included in the cost frontier function) is estimated with SVKZ model. The results for the four cases of specifications (raw output and casemix adjusted output with linear form and log-log transformation) are reported in Table B-1.

The estimated results show that the mean inefficiency scores of the industry are similar under the two approaches: parametric vs. semi-parametric. This is especially for the log-log transformation, where the heteroskedasticity of the data is partially accounted for. This is an indication of the robustness of the results.

The estimated frontier under two approaches are also examined. Figure B-1 to Figure B-4 present the estimated parametric and semi-parametric frontiers for each year. The parametric and semi-parametric frontiers are close together, especially for the case of casemix adjusted output, where the two estimated frontiers nearly overlap. This is another indication of the robustness of the results.

Based on the comparison, the stochastic frontier models applied so far are robust to the parametric assumption, especially for the log-log transformation with casemix adjusted output, where both average inefficiency and the shape of the frontier estimated from stochastic frontier models are very close to those estimated using the non-parametric model.

Table B-1. Estimated mean efficiency scores with the ALS77 and SVKZ approaches

Model	Raw output (linear form)	Casemix adjusted output (linear form)	Raw output (log-log transformation)	Casemix adjusted output (log-log transformation)
ALS77	20.38%	20.15%	14.91%	14.52%
SVKZ	30.10%	28.85%	17.79%	13.83%

ALS77: Aigner, Lovell and Schmidt (1977); SVKZ: Simar, Van Keilegom and Zelenyuk (2017). See Section 4.1 in Technical Supplement Report 2

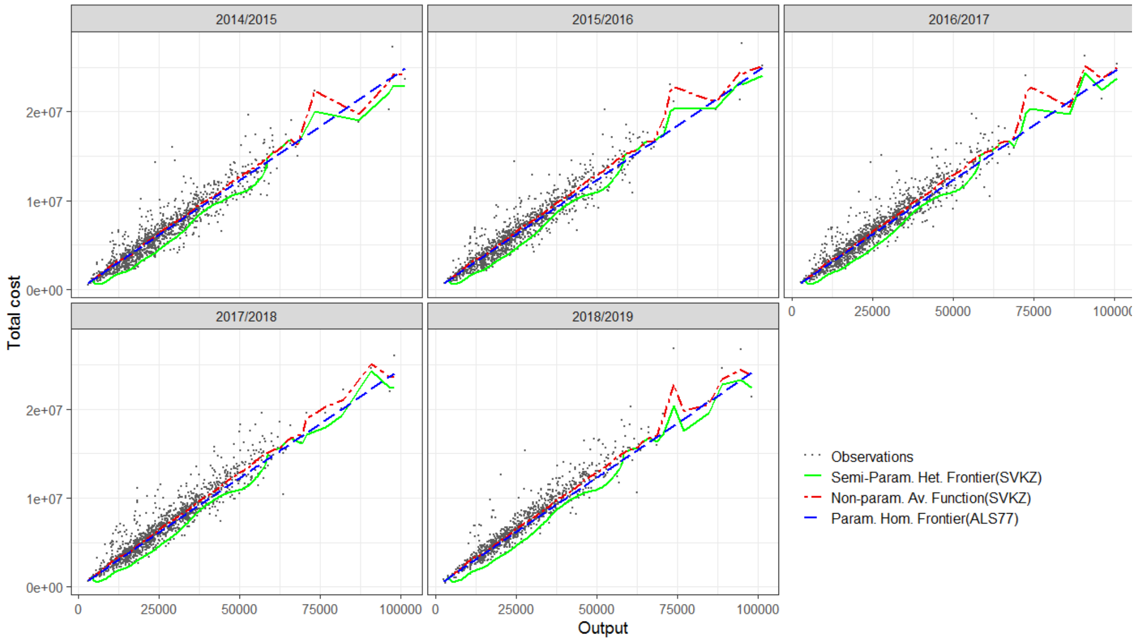

Figure B-1. Estimated frontiers: SKVZ vs. ALS77 for linear form, raw output

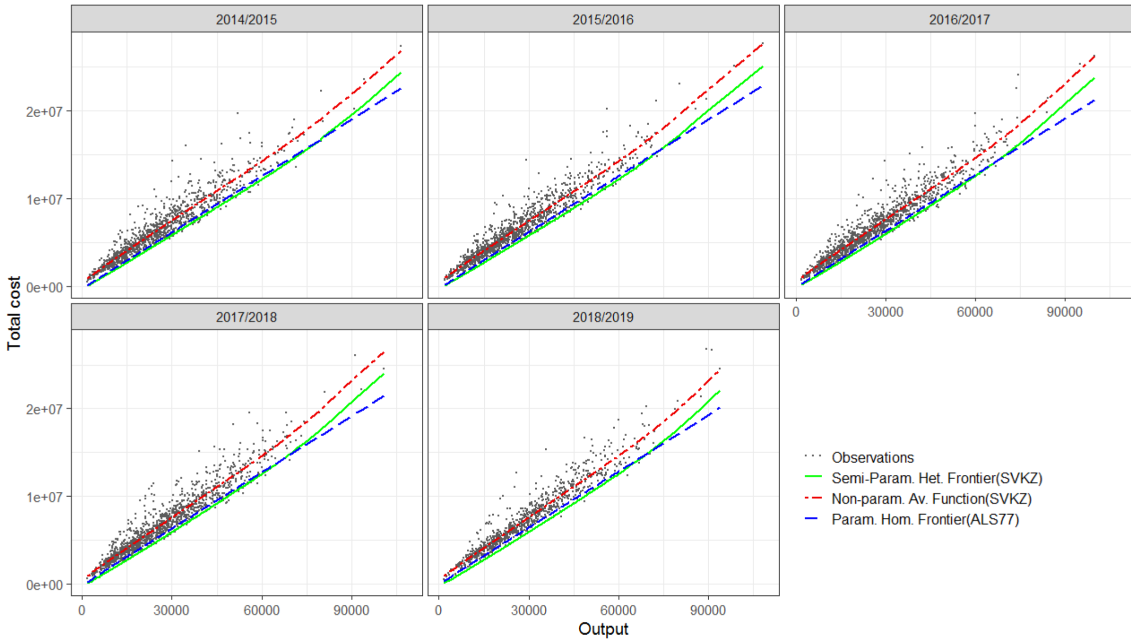

Figure B-2. Estimated frontiers: SKVZ vs. ALS77 for linear form, casemix adjusted output

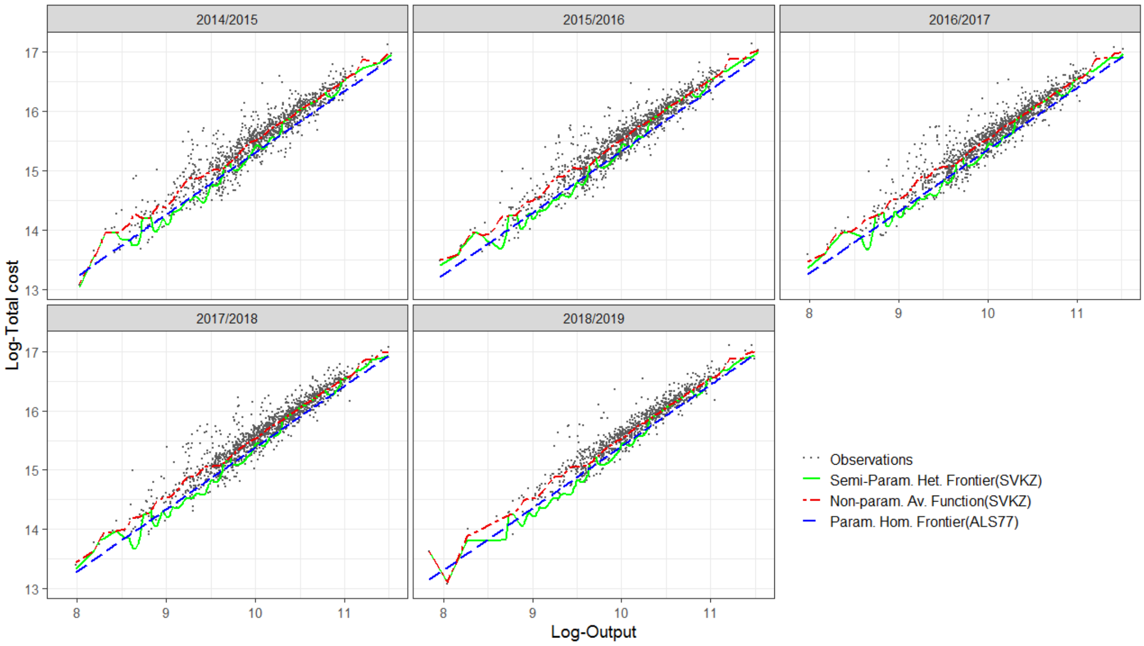

Figure B-3. Estimated frontiers: SKVZ vs. ALS77 for log-log transformation, raw output

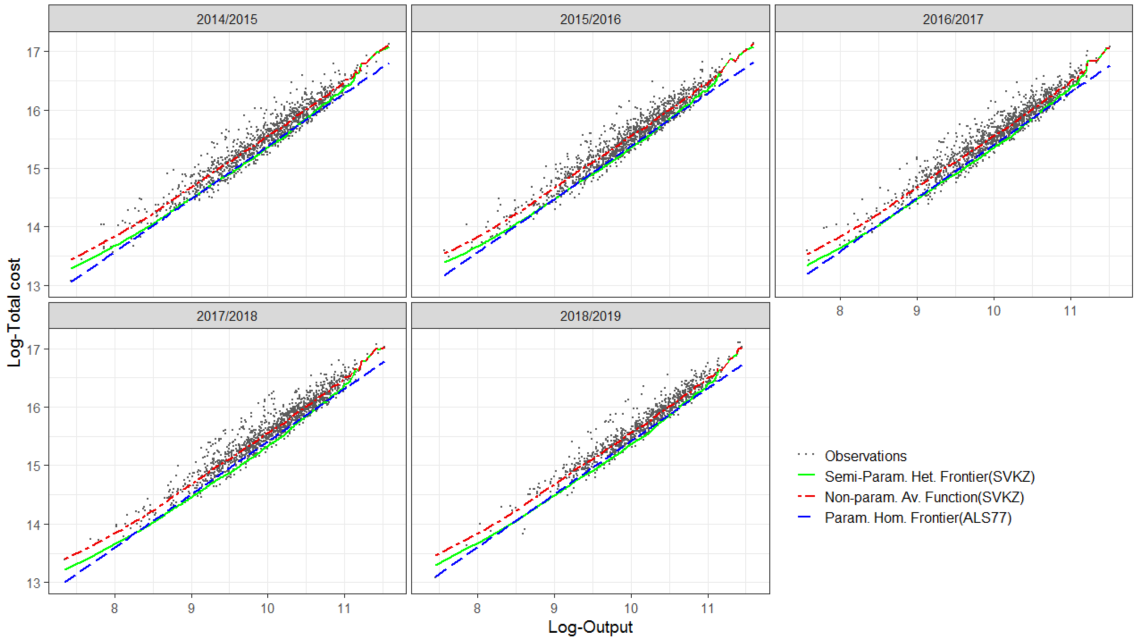

Figure B-4. Estimated frontiers: SKVZ vs. ALS77 for log-log transformation, casemix adjusted output

Appendix C
Additional results from the stochastic frontier analysis

A large number of frontier models were estimated to check for robustness, and to further understand the behaviours of the residential aged care sector. In this appendix, Table C-1 to Table C-8 provides additional estimated results where facility types were included in the cost frontier. As noted in TSR2, it is important to realise that the differences in efficiency estimated between for-profit, not-for-profit and government facilities may (and perhaps are likely to) reflect differences in quality achieved by the different ownership types which have not been able to be distinguished within the 3 levels of the composite quality index.

C-1 Estimation for linear function, raw output with facility types

Table C-1. Frontier estimation for linear functional form, raw output (occupied bed days) with facility types included in the cost frontier function

Total cost	ALS77	PL81	SS84
Frontier			
Y (occupied bed days)	244.29*** (1.24)	211.77*** (2.03)	196.23*** (3.24)
Financial year 2014/15	-569679.9*** (47569.52)	-491558.6*** (23228.52)	-502429.8*** (22768.03)
Financial year 2015/16	-440679.8*** (46970.65)	-373507.2*** (22904.1)	-380498.8*** (22394.87)
Financial year 2016/17	-314680.5*** (46136.89)	-244364.8*** (22237.74)	-249649*** (21678.12)
Financial year 2017/18	-176349.8*** (46016.78)	-134957.1*** (22034.87)	-141006.6*** (21463.1)
For-profit	-63095.7 (102510.1)	-343296.2** (141039.6)	-182465.9 (303405)
Government	7866.77 (142777.1)	-196224.3 (246736.5)	-121490.8 (480997.7)
Y x For-profit	11.45*** (3.29)	26.9*** (4.5)	43.33*** (7.49)
Y x Government	24.63*** (8.38)	22.88 (13.91)	30.28 (24.16)
Constant	-611031.8*** (47157.32)	-305409.6*** (53745.48)	1666757*** (93960.89)
$ln(\sigma_u^2)$ ***(Inefficiency)***			
Constant	28.42*** (0.04)	29.06*** (0.05)	
$ln(\sigma_v^2)$ ***(Residuals)***			
Constant	26.89*** (0.05)	26.21*** (0.02)	

ALS77: Aigner, Lovell and Schmidt (1977); PL81: Pitt and Lee (1981); SS84: Schmidt and Sickles (1984). See Section 4.1 in Technical Supplement Report 2.
The base case is: financial year 2018/19, not-for-profit provider.
Standard errors in parentheses
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table C-2. Estimated inefficiency for linear functional form, raw output with facility types included in the cost frontier function

	Model	Observations	Mean	Standard deviation	Minimum	Maximum
Inefficiency	ALS77	5,711	1,147,635	775,353	104,173	8,319,854
	PL81	5,711	1,643,665	1,159,376	28,843	9,643,336
	SS84	5,711	3,356,168	1,333,869	0	11,700,000
Inefficiency to total cost	ALS77	5,711	0.20	0.14	0.01	1.86
	PL81	5,711	0.25	0.14	0.00	1.09
	SS84	5,711	0.59	0.29	0.00	3.81

ALS77: Aigner, Lovell and Schmidt (1977); PL81: Pitt and Lee (1981); SS84: Schmidt and Sickles (1984). See Section 4.1 in Technical Supplement Report 2.

C-2 Estimation for linear function, casemix adjusted output with facility types

Table C-3. Frontier estimation for linear functional form, casemix adjusted output with facility types included in the cost frontier function

Total cost	ALS77	CFG95	SS84	G05
Frontier				
Y* (casemix adjusted occupied bed days)	214.19*** (1.07)	215.03*** (1.07)	142.42*** (2.82)	142.42*** (2.44)
Financial year 2014/15	-269068.4*** (38088.52)	-271283.2*** (37973.55)	-339597.7*** (25294.15)	-339597.7*** (21945.61)
Financial year 2015/16	-266533*** (37549.39)	-269478.9*** (37402.82)	-300643.7*** (24500.37)	-300643.7*** (21256.92)
Financial year 2016/17	-139224.5*** (37029.21)	-141610.9*** (36868.71)	-160389.1*** (23682.31)	-160389.1*** (20547.16)
Financial year 2017/18	-82655.65** (36725.63)	-89073.53** (36629.25)	-96332.36*** (23351.82)	-96332.36*** (20260.42)
For-profit	-199464.2** (82801.46)	-271254.5*** (86762.12)	32146.45 (307408.2)	32146.45 (266712.3)
Government	-335198.8*** (107043.6)	-310757.5*** (107717.7)	223537.7 (457135.7)	223537.7 (396618.3)
Y* x For-profit	0.44 (2.48)	1.44 (2.55)	48.1*** (6.85)	48.1*** (5.95)
Y* x Government	30.63*** (6.02)	28.76*** (6.13)	6.64 (21.76)	6.64 (18.88)
Constant	4699.83 (37566.54)	1599.95 (37312.68)	2942566*** (87220.73)	642750.2 (931742.4)
$ln(\sigma_u^2)$ ***(Inefficiency)***				
Constant	28.56*** (0.03)	29.95*** (0.2)		0.96 (1409550)
Occupancy rate		-0.02*** (0.002)		
$ln(\sigma_v^2)$ ***(Residuals)***				
Constant	25.77*** (0.07)	25.82*** (0.07)		25.82*** (0.07)

ALS77: Aigner, Lovell and Schmidt (1977); CFG95: Caudill, Ford, and Gropper (1995); SS84: Schmidt and Sickles (1984); G05: Greene (2005). See Section 4.1 in Technical Supplement Report 2.
The base case is: financial year 2018/19, not-for-profit provider.
Standard errors in parentheses
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table C-4. Estimated inefficiency for linear functional form, casemix adjusted output with facility types included in the cost frontier function

	Model	Observations	Mean	Standard deviation	Minimum	Maximum
Inefficiency	ALS77	5,711	1,216,021	969,871	95,037	8,443,118
	CFG95	5,711	1,205,084	960,863	95,903	8,575,727
	SS84	5,711	2,814,931	1,680,900	0	12,600,000
	G05	5,711	1	0	1	1
Inefficiency to total cost	ALS77	5,711	0.20	0.13	0.01	1.00
	CFG95	5,711	0.20	0.13	0.01	1.01
	SS84	5,711	0.43	0.13	0.00	1.80
	G05	5,711	0.00	0.00	0.00	0.00

ALS77: Aigner, Lovell and Schmidt (1977); CFG95: Caudill, Ford, and Gropper (1995); SS84: Schmidt and Sickles (1984); G05: Greene (2005). See Section 4.1 in Technical Supplement Report 2.

C-3 Estimation for log function, raw output with facility types

Table C-5. Frontier estimation for logarithmic functional form, raw output with facility types included in the cost frontier function

LnTC (log of total cost)	ALS77	CFG95	SS84	PL81	K90	CSW14
Frontier						
lnY (log of occupied bed days)	1.05*** (0.005)	1.05*** (0.005)	0.71*** (0.013)	0.98*** (0.009)	0.96*** (0.008)	0.74*** (0.013)
Financial year 2014/15	-0.1*** (0.008)	-0.1*** (0.008)	-0.09*** (0.003)	-0.08*** (0.004)	-0.18*** (0.008)	-0.09*** (0.003)
Financial year 2015/16	-0.07*** (0.008)	-0.07*** (0.008)	-0.06*** (0.003)	-0.06*** (0.004)	-0.14*** (0.007)	-0.05*** (0.003)
Financial year 2016/17	-0.05*** (0.008)	-0.05*** (0.008)	-0.04*** (0.003)	-0.04*** (0.003)	-0.1*** (0.007)	-0.03*** (0.003)
Financial year 2017/18	-0.02*** (0.008)	-0.03*** (0.008)	-0.02*** (0.003)	-0.02*** (0.003)	-0.05*** (0.005)	-0.01*** (0.003)
For-profit	0.36** (0.154)	0.08 (0.161)	-1.05*** (0.27)	-0.21 (0.206)	-0.1 (0.203)	-1.1*** (0.283)
Government	0.74*** (0.229)	0.86*** (0.229)	-2.91*** (0.444)	-0.25 (0.327)	-0.65* (0.368)	-3.27*** (0.355)
lnY x For-profit	-0.03** (0.015)	-0.01 (0.016)	0.12*** (0.026)	0.03 (0.021)	0.01 (0.021)	0.12*** (0.028)
lnY x Government	-0.07*** (0.024)	-0.08*** (0.024)	0.32*** (0.047)	0.03 (0.034)	0.07* (0.039)	0.36*** (0.038)
Constant	4.89*** (0.053)	4.85*** (0.052)	8.42*** (0.131)	5.46*** (0.082)	5.68*** (0.076)	
$ln(\sigma_u^2)$ ***(Inefficiency)***						
Constant	-3.11*** (0.053)	-0.61*** (0.229)		-1.94*** (0.051)	-1.59*** (0.071)	
Occupancy rate		-0.03*** (0.003)				
$ln(\sigma_v^2)$ ***(Residuals)***						
Constant	-4.06*** (0.042)	-4.06*** (0.042)		-5.18*** (0.023)	-5.25*** (0.023)	
G(t)						
t					0.03 (0.037)	
t^2					0.01** (0.007)	

ALS77: Aigner, Lovell and Schmidt (1977); CFG95: Caudill, Ford, and Gropper (1995); SS84: Schmidt and Sickles (1984); PL81: Pitt and Lee (1981); K90: Kumbhakar (1990); CSW14: Chen, Schmidt, and Wang (2014). See Section 4.1 in Technical Supplement Report 2.
G(t): time-varying component; t: time.
The base case is: financial year 2018/19, not-for-profit provider.
Standard errors in parentheses
* p < 0.10; ** p < 0.05; *** p < 0.01

Table C-6. Estimated efficiency scores for logarithmic functional form, raw output with facility types included in the cost frontier function

Model	Observations	Mean	Standard deviation	Minimum	Maximum
ALS77	5,711	85.40%	7.26%	44.10%	96.96%
CFG95	5,711	85.68%	7.39%	38.19%	96.86%
SS84	5,711	36.54%	9.08%	14.28%	100.00%
PL81	5,711	72.62%	11.56%	29.97%	99.49%
K90	5,711	71.65%	11.80%	27.20%	99.50%
CSW14	5,711	91.20%	4.58%	59.27%	99.95%

ALS77: Aigner, Lovell and Schmidt (1977); CFG95: Caudill, Ford, and Gropper (1995); SS84: Schmidt and Sickles (1984); PL81: Pitt and Lee (1981); K90: Kumbhakar (1990); CSW14: Chen, Schmidt, and Wang (2014). See Section 4.1 in Technical Supplement Report 2.

C-4 Estimation for log function, casemix adjusted output with facility types

Table C-7. Frontier estimation for logarithmic functional form, casemix adjusted output with facility types included in the cost frontier function

LnTC (log of total cost)	ALS77	CFG95	SS84	PL81	K90	KW05	CSW14
Frontier							
lnY* (log of casemix adjusted occupied bed days)	0.91***	0.91***	0.52***	0.84***	0.85***	0.85***	0.52***
	(0.004)	(0.003)	(0.011)	(0.006)	(0.006)	(0.006)	(0.011)
Financial year 2014/15	-0.03***	-0.03***	-0.05***	-0.02***	-0.06***	-0.06***	-0.05***
	(0.006)	(0.006)	(0.004)	(0.004)	(0.007)	(0.007)	(0.004)
Financial year 2015/16	-0.03***	-0.03***	-0.04***	-0.03***	-0.07***	-0.06***	-0.04***
	(0.006)	(0.006)	(0.003)	(0.004)	(0.007)	(0.005)	(0.003)
Financial year 2016/17	-0.01*	-0.01*	-0.02***	-0.01*	-0.04***	-0.02***	-0.02***
	(0.006)	(0.006)	(0.003)	(0.004)	(0.007)	(0.004)	(0.003)
Financial year 2017/18	-0.004	-0.01	-0.01***	-0.004	-0.03***	-0.01***	-0.01***
	(0.006)	(0.006)	(0.003)	(0.004)	(0.005)	(0.004)	(0.003)
For-profit	-0.25**	-0.44***	-1.55***	-0.21	-0.17	-0.17	-1.63***
	(0.113)	(0.119)	(0.243)	(0.148)	(0.148)	(0.147)	(0.24)
Government	-0.14	-0.14	-1.11***	0.15	0.08	0.1	-1.2***
	(0.169)	(0.169)	(0.34)	(0.235)	(0.239)	(0.234)	(0.336)
lnY* x For-profit	0.02*	0.04***	0.17***	0.02	0.02	0.02	0.18***
	(0.011)	(0.012)	(0.024)	(0.015)	(0.015)	(0.015)	(0.024)
lnY* x Government	0.02	0.02	0.14***	-0.02	-0.01	-0.01	0.15***
	(0.018)	(0.018)	(0.038)	(0.025)	(0.026)	(0.025)	(0.037)
Constant	6.36***	6.35***	10.38***	6.93***	6.92***	6.91***	
	(0.037)	(0.036)	(0.108)	(0.056)	(0.056)	(0.056)	
$ln(\sigma_u^2)$ ***(Inefficiency)***							
Constant	-3.15***	-1.07***		-2.83***	-2.89***	-2.71***	
	(0.04)	(0.2)		(0.052)	(0.102)	(0.054)	
Occupancy rate		-0.02***					
		(0.002)					
$ln(\sigma_v^2)$ ***(Residuals)***							
Constant	-4.89***	-4.86***		-4.97***	-4.98***	-4.97***	
	(0.057)	(0.056)		(0.022)	(0.022)	(0.022)	
G(t)							
t					-0.18**		
					(0.075)		
t^2					0.05***		
					(0.013)		
$t - \underline{t}$						-0.05***	
						(0.007)	

ALS77: Aigner, Lovell and Schmidt (1977); CFG95: Caudill, Ford, and Gropper (1995); SS84: Schmidt and Sickles (1984); PL81: Pitt and Lee (1981); K90: Kumbhakar (1990); KW05: Kumbhakar and Wang (2005); CSW14: Chen, Schmidt, and Wang (2014). See Section 4.1 in Technical Supplement Report 2.
G(t): time-varying component; t: time; $\underline{t}$: the beginning period of sample.
The base case is: financial year 2018/19, not-for-profit provider.
Standard errors in parentheses
* p < 0.10; ** p < 0.05; *** p < 0.01

Table C-8. Estimated efficiency scores for logarithmic functional form, casemix adjusted output with facility types included in the cost frontier function

Model	Observations	Mean	Standard deviation	Minimum	Maximum
ALS77	5,711	85.56%	8.41%	48.80%	97.98%
CFG95	5,711	85.85%	8.35%	47.09%	97.96%
SS84	5,711	33.41%	9.49%	14.43%	100.00%
PL81	5,711	82.74%	9.86%	43.43%	99.16%
K90	5,711	82.89%	9.88%	42.52%	99.26%
KW05	5,711	82.92%	9.86%	41.47%	99.23%
CSW14	5,711	92.89%	3.39%	65.25%	99.58%

ALS77: Aigner, Lovell and Schmidt (1977); CFG95: Caudill, Ford, and Gropper (1995); SS84: Schmidt and Sickles (1984); PL81: Pitt and Lee (1981); K90: Kumbhakar (1990); KW05: Kumbhakar and Wang (2005); CSW14: Chen, Schmidt, and Wang (2014). See Section 4.1 in Technical Supplement Report 2.

References

Aigner, D., Lovell, C. A. K., and Schmidt, P. (1977), “Formulation and Estimation of Stochastic Frontier Production Function Models,” *Journal of Econometrics*, 6, 21–37.

Caudill, S. B., Ford, J. M. and Gropper, D. M. (1995), “Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroscedasticity,” *Journal of Business and Economic Statistics*, 13, 105–11.

Chen, Y.-Y., Schmidt, P., and Wang, H.-J. (2014), “Consistent Estimation of the Fixed Effects Stochastic Frontier Model,” *Journal of Econometrics*, 18(2), 65–76.

Greene, W. H. (2005), “Fixed and Random Effects in Stochastic Frontier Models,” *Journal of Productivity Analysis*, 23, 7–32.

Kumbhakar, S. C. (1990), “Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency,” *Journal of Econometrics*, 46, 201–11.

Kumbhakar, S. C., and Wang, H.-J. (2005), “Estimation of Growth Convergence Using a Stochastic Production Frontier Approach,” *Economics Letters*, 88, 300–5.

Pitt, M. M., and Lee, L.-F. (1981), “The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry,” *Journal of Development Economics*, 9, 43–64.

Schmidt, P., and Sickles, R. C. (1984), “Production Frontiers and Panel Data,” *Journal of Business and Economic Statistics*, 2, 367–74.

Sickles, R. C., and Zelenyuk, V. (2019), *Measurement of Productivity and Efficiency*, Cambridge University Press.

Simar, L., Van Keilegom, I., and Zelenyuk, V. (2017), "Nonparametric least squares methods for stochastic frontier models," *Journal of Productivity Analysis*, *47*(3), 189-204.

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

CREATE CHANGE

The University of Queensland

CRICOS Provider Number 00025B