**Litterature Review**

3.1.2

Analysis on Citations

The citation scores of each submission were provided by Scopus after request of the REF for

input, but not for all the UOAs [REF, 2014a]. According to the panel criteria and working

methods, as supplied by the panel, [REF, 2012c] the citation counts for Panel A, Panel B and

Panel C where used as a supplementary tool as part of the indication of academic significance of

the published work. Panel D on the other hand clearly specified in the assessment criteria [REF,

2012c], that they would neither receive or make use of citation data to assess the submitted

work.Panels A and B appeared to have given more emphasis on the citation count than Panel

C, with Panel B even utilising additional citation data from Google Scholar as an indicator.

Panel B criteria [REF, 2012a] have specified that Google Scholar was accessed in a systematic

way to observe if there was work highly cited outside of Scopus, which raises some issues. It

does not specify if the Google Scholar citation data was retrieved during the same period for

all papers and to what extent was this data used. Nevertheless , what one can interpret from

the published criteria from these panels, is that the citation data could partly be responsible

for the output assessment, except for Panel D.

Previous work on how citations could have influence the outcome of the REF 2014 scores have

been performed by Alan Dix [Dix, 2015a] and Morris Sloman [Sloman, 2014]. Morris Sloaman,

who was also the deputy chain in sub-panel UOA 11 [Sloman, 2014]. He was responsible for

assessing the work submitted under the Computer Science and Informatics UOA. Therefore, the

author attempted to analyse the REF output results by using citation-based metrics. Citation

data from Google Scholar, as well as from Scopus were used to observe the percentage of REF

ranked output against the citation data from both of the sources. Interpreting Morris Sloman’s

results,it is observed that there is an overall correlation between the REF ratings and citations.

Alan Dix, who was a member of the sub-panel responsible for assessing the Computer Science

and Informatics UOA, also examined the work done by Morris Sloman [Sloman, 2014] and by

others for the sub-areas in this UOA and found a systematic pattern. The pattern was that in

several analyses performed the theoretical sub-areas of Computer Science and Informatics were

benefiting over the practical ones [Dix, 2015b].

In Alan Dix’s work Citations and Sub-Area Bias in the UK Research Assessment Process

[Dix, 2015a], he has also produced work to assess the fairness of the REF process again explicitly

in the Unit of Assessment he was responsible for assessing. The author’s work focused mainly

on assessing whether the citation counts have influenced the outcome of the scores. He also

examined further the use of citation data to uncover if there was any bias within the sub-areas

8of Computer-Science and Informatics. The author used publicly available data for this analysis

including raw Scopus data for the period of the process, the normalised Scopus data submitted

to the REF and Google Scholar data. The Computer Science and Informatics UOA had the

citation data supplied by Scopus. The results suggested that there was bias between sub-areas

leading several universities to benefiting more than others and possibly also introducing gender-

bias [Dix, 2015a].

3.1.3

Analysis on Income / Journal Ranking

Although Morris Slomans work [Sloman, 2014] mainly focused on citation based bias, the author

also researched other metrics that could correlate with the REF output ratings. Examining

Income per FTE against the REF scores, the author uncovered that there is strong correlation

between the two, which can suggest that a higher research income per staff can produce higher

REF scores. Additionally, the author used journal ranking data from Scopus and Thomson as

metrics to assess the REF 4\* percentages. It appears that there is no correlation between the

journal ratings and the percentage 4\*.

**Methodology**

4.2.1

Excellence in Research for Australia (ERA)

Similar to the REF, the ERA is a process initialised by the Australian Government to assess the

quality of work of the higher education institutions in Australia [Australian Research Council,

2014]. When ERA was first conducted in 2010, as part of the process, a normalised spreadsheet

was produced which ranked international journals based on inputs from expert, academic and

public. The spreadsheet was then used as part of the assessment to evaluate the work of the

higher institutions in Australia. The analysis on the results further on proceeded with every

caution, since there is a time gap between the 2010 publication of the journal rankings and the

2014 REF process.

4.2.2

Higher Education Statistics Agency (HESA)

HESA was legally formed in 1993 as a not-for-profit private entity and has a vision of becoming

HEIs related analytical powerhouse in UK. This organization works with HEIs in supporting

their reporting of data under statutory requirement by UK Government. During REF 2014,

HESA has provided contextual data related to estimates of total Full-Time Equivalent (FTE)

expected from HEIs as part of REF 2014 analysis requirement [HESA, 2014].

4.2.3

SCimago Journal and Country Rank

The SCImago Journal & Country Rank is a publicly available portal that includes the jour-

nals and country scientific indicators developed from the information contained in the Scopus

database [Scimago Journal and Country Rank, 2017]. The 2014 dataset containing international

journals with the SJR score that SCImago uses to rank journals, which is an indicator based

on citation data from Scopus, was used for this analysis.

4.6

Feature 2 - Citations Count

Alan Dix’s [Dix, 2015a] and Morris Sloman’s [Sloman, 2014] work focused only on UOA 11 for

the citation analysis, and not any of the other UOAs. Additionally the outcome from reading

through assessment criteria specified by the Main Panels[REF, 2012c], is that the citations

should have some effect on the output rating but should not be a deciding factor. Main Panel D

is the exception though as they clearly specified that they would not request or use any citation

data to assess submitted work [REF, 2012c].

This analysis therefore is focused on identifying if there is any correlation between citations

and the output ratings by the REF for all UOAs. The results will then be used to assess the

difference in the extent of use of the citations as an indicator for the performance of work in

different sub-panels and UOAs.

From the UOA’s selected the following had citation data supplied by Scopus after request

by the REF:

• 2 - Public Health, Health Services and Primary Care (Panel A)

• 5 - Biological Sciences (Panel A)

• 9 - Physics (Panel B)

• 11 - Computer Science & Informatics (Panel B)

• 18 - Economics and Econometrics (Panel C)

The analysis was therefore performed only on the UOAs listed above, due to the confirmed

source of the citation data [REF, 2014a], using the data provided by the REF [REF, 2014b].

The original data contained the citation counts as integers in almost all UOAs, including 0, for

each submission and in certain cases data was not available and the citation count was given

as blank. To account for the unavailable data, instead of comparing the output score against

the total citations per university, the total citations were divided by the number of submissions

in each university to obtain a normalised value. UOA 11 was the only UOA were the citation

data was supplied as string numbers and where therefore converted to integers.

The citations per submission for each university were then compared against the percentage

of output ranked 4\* and the combination of 2\* and 1\* papers in each university. The Pearsons

method was used to return the correlation of the two variables. The reason why the combination

2\* and 1\* output was used, is to confirm that the inverse relationship that is expected from the

4\* output is true.

4.7

Feature 3 - Research Income

In Morris Sloman’s work [Sloman, 2014], income by staff is used to check whether there was any

correlation between that and the REF outputs, in UOA 11. This analysis aims at examining

whether there is any correlation between income per submission for each university in each UOA

and the REF output. Through the results it is expected to observe if generally higher research

income, regardless of the income source, implies higher quality research and thus a higher REF

output.

The original data obtained from the REF website for each UOA, contained the research

income by source for each university for each year during the period of the process. To obtain

14the income per submission per university, the total income for each university, for all sources

and years was calculated. The incomes were supplied as integers. This number was then

divided by the total number of submissions in each university. The income per submission

for each university was then compared against the percentage of 4\* and a combination of the

percentages of 2\* and 1\* outputs. It is expected that universities with higher income per

submission, regardless of the income source, will have a higher percentage of 4\* output and a

lower percentage of 2\* and 1\* outputs.

4.8

Feature 4 - Journal Ranking

Morris Sloman’s [Sloman, 2014] work did not show any correlation between journal ranking and

the REF output. The main aim of this assessment was to examine the hypothesis is that the

journals and their reputation, or ranking, in which the research work was submitted to affects

the 4\* output score of the paper. The ERA 2010 journal ranking list [Australian Research

Council, 2014], was used as the base dataset for the analysis on all the UOAs and the SCImago

Journal Ranking list for the year 2014 for a complementary analysis on the results from the

base dataset [Scimago Journal and Country Rank, 2017]. The SCImago Journal Ranking List

was discovered later on in the analysis of the results, and therefore due to time limitations it

was used as a supplementary dataset for confirming the results of the base dataset.

4.8.1

2010 ERA Journal Ranking List

The ERA 2010 journal ranking list ranks journals by A\*, A, B, C and Not Ranked, with A\*

being the highest rated and C the lowest rating. The original data from the REF contains (per

UOA) contains the submission, volume title and the ISSN code, which is a unique number for

each journal.

Since there is a four year gap between the publication of the REF and the 2010 ERA, it is

assumed that the journal ranks remained unaffected. Additionally due to the previous fact and

the fact that not all work submitted in the REF was published in journals, it was expected that

there would be some data loss when matching the two data sets.

The 2010 ERA journal ranking list and the REF data for each UOA were initially matched

by volume titles, with most UOAs having half of the data matched. To match by volume titles,

the text was normalised by lower-casing, and punctuation and any symbols were removed. This

raised the issue though that there might be a mismatch between similarly named journals. To

avoid this it was decided that it would be better if the datasets were matched by the ISSN

numbers.

This returned a matched dataframe which contained the university, submission, volume title

and the 2010 ERA rank. To be able to quantify the effect the journal ranks have on the output,

the journal rank scores were given a representative score. The scores were converted as follows:

• A\* : 4

• A:3

• B:2

• C:1

• Not Ranked : 0

The idea behind the scores is that a paper published in an A\* journal will have a larger

gravity in affecting the output 4\* percentage of the university, C will have the least and Not

Ranked will have none. Different weighting schemes were trialled on several UOAs before

settling with the scores above. Different scores produced better correlation in some cases, for

15example in the Physics UOA, the scores A\*: 1, A : 1, B : -2.5, C: -5 and Not Ranked:0, returned

a correlation of 0.61. This was better but the pattern was the same and the overall outcome

was not largely affected. After converting the 2010 ERA ranks to their scores, the scores were

summed up for each university and divided by the number of submissions to find the average

journal rank per university. The average journal rank per university was then compared against

the 4\* output.

4.8.2

SCImago Journal Ranking List

The SCImago Journal Ranking List ranks journals by an indicator, the SJR score, which is

based on data from the Scopus data created by SCImago. This dataset was used only in

strongy correlated UOAs from the 2010 ERA Journal Ranking List. Unlike with the 2010 ERA

dataset, the SCImago dataset and the REF dataset were matched by the name of the journal

title rather than the ISSN. Matching by ISSN, returned very low matched percentages, whilst

matching by journal name after normalisation resulted in higher matching percentages.

Following the data matching by name the new dataframe for each UOA contained the

university, submission, journal title and the SJR . The SJR scores where then summed up for

each university and divided by the number of submissions to find the average SJR score per

submission per university in each UOA. This value was then compared against the output 4\*

of each university.

**Results**

5.2

Feature 2 - Citations Count

The R values from the Pearson test performed on the data can be seen in Table 4. The R-values

returned indicate the correlation between average citations per submission for each university

in each UOA and the REF output listed in the table.

21Table 4: Pearson correlation R-values between average citations per submission vs. Output 4\*

/ Output 2\* + 1\*. \*\* refers to strong Pearson correlation; \* refers to moderate correlation.

To illustrate the two best correlated UOAs, Figure 5 can be seen below. From the figures

it appears that there is a linear relationship between the two data and this is supported by the

R-values returned. When comparing against the combined 2\* and 1\* outputs the relationship

in the Economics and Econometrics UOA appears to follow a linear-to-quadratic pattern.

Figure 5: Average citations per submission vs. REF output percentages 4\* and (2\* +1\*) for UOAs 2

and 11.

The results suggest that there is large variance between the effect of the citation numbers

and output scores in each UOA. Even comparing the correlation results between UOAs of the

same panel, we can see some variances, with the obvious example being Main Panel A. What this

can suggest is that although there were guidelines on the assessment criteria, when considering

citations, published by each Main Panel [REF, 2012c], several sub-panels used the citation data

more than others.

In UOA 18, which is the UOA with the strongest correlations on both outputs compared,

the results can imply that the sub-panel responsible for assessing the research work did not

adhere to the degree described by Main Panel C in the assessment criteria [REF, 2012c].

5.3

Feature 3 - Research Income

The R values from the Pearson test performed on the data can be seen in Table 5. The R-values

returned indicate the correlation between average income per submission for each university in

22each UOA and the REF output listed in the table.

Table 5: Pearson correlation R-values between average income per submission vs. Output 4\* /

Output 2\* + 1\*. \*\* refers to very strong Pearson correlation; \* refers to strong correlation.

The results indicate that there is a large variance on the effect of income on the REF output

result between UOAs. This variance is also observed between UOAs of the same Main Panel,

with the example of Main Panel A, similarly as in the citations. The larger number of strongly

correlated UOA’s can suggest that the research quality is largely dependent, in these areas, and

heavily influenced by the research income per submission.

5.4

5.4.1

Feature 4 - Journal Rankings

Results using the ERA 2010 Journal Ranking List

The R values from the Pearson test performed on the data can be seen in Table 6, along with

the percentage of matched data in each UOA. The R-values returned indicate the correlation

between average journal ranking for each university in each UOA and the REF output listed in

the table.

Table 6: Pearson correlation R-values between average income per submission vs. Output 4\*.

Table also includes matched data in percentage for each UOA. \*\* refers to very strong Pearson

correlation; \* refers to strong correlation.

The results from several UOAs returned a very strong positive correlation, therefore possibly

implying that the journal rank does affect the output score of a paper. In an attempt to further

analyse the results the Computer Science and Informatics (UOA 11) and the Economics and

Econometrics (UOA 18) were selected for further analysis. The figures below show the plotted

results for both of the UOAs, with the analysis on each graph below them.

23Figure 6: UOA 11 - Average journal rank vs. Percentage of 4\* Output. The annotations

represent the UKPRN code for the universities. Each data point is a university.

What is visible from Figure 6 is that almost all of the universities with up to approximately

2.3 in average journal rank have 0% of 4\* output rated papers. After that threshold the data

appears to follow a linear pattern, with universities having a higher average journal rank having

also a higher percentage of 4\* output. What is interesting are the universities that appear

not to follow the pattern, or are no close to the rest of the data. Specifically universities with

UKPRN 10000886 and 10007161 have a high average journal rank but a very low percentage of

4\* output papers, when comparing with universities with similar average journal rank. What

was observed is that these universities have significantly less submissions than the universities

with an average rank close to these two.

24Figure 7: UOA 18 - Average journal rank vs. Percentage of 4\* Output. The annotations

represent the UKPRN code for the universities. Each data point is a university.

By selecting the university with the UKPRN 10005553 from Figure 7 and comparing to the

universities with a higher 4\* output but similar journal rank, the same pattern is observed.

University 10005553 has significantly less submissions than the universities with a close average

journal rank.

5.4.2

Further analysis using the SCImago 2014 Journal Ranking List

To further assess the most strongly correlated UOAs from Panels A, B and C to check see if

the hypothesis that journal rankings could possibly affect the REF output results still holds

valid. Journal ranking scores from the SCImago Journal & Country Rank [Scimago Journal and

Country Rank, 2017] for the year 2014 were used. Table 7 summarises the R-values returned with

the Pearson test, between the average SJR score per institution and the output 4\* percentage.

Table 7: R-values of correlation between average SJR score per institution vs. Output 4\* for

UOAs 2, 11 and 18.

The results produced with the SCImago dataset agree with the results produced with the

2010 ERA Journal Ranking List. The correlation results of the three UOAs still indicate that

there is a strong correlation between the journal ranking in any sort of form against the 4\*

output of the REF. In this case the results even show that in UOA 18 the two variables are

very strongly correlated. The plot for UOA 18 can be seen below in Figure 8.

25Figure 8: UOA 18 - Average SJR score per submission vs. Percentage of 4\* Output. Each data

point is an institution

**Discussion**

6.2

Feature 2 - Citations

The results have indicated variances in correlation values between the UOA’s within the same

Main Panel, but also generally. How this can be interpreted is that the sub-panels responsible

for the UOAs with higher correlation values, may have possibly used to a larger extent the

citations as an indicator of the quality of submitted work.

6.3

Feature 3 - Research Income

The strong correlation results in several UOAs can suggest that in those UOAs, the quality of

work is largely affected by the research income, irrespective of the source. The variance between

correlation results in the UOAs is still apparent.

In UOAs where the results indicate strong correlations, it is possible that higher income per

submission is an indicator of higher quality work. There is also the possibility that although the

assessing members were trying to be objective, there was some institutional bias in the strongly

correlated UOAs. The assumption is that papers from higher research income institutions were

assessed more leniently that lower research income institutions.

326.4

Feature 4 - Journal Ranking

By using the ERA 2010 Journal Ranking List [Australian Research Council, 2014], the results

show strong correlations in several UOAs. In particular there is a repetitive pattern with the

Research Income feature. As it appears the three UOAs with the strongest correlation results

in both cases are UOAs 5, 11, 18 and the UOA with the weakest correlation in both cases is

UOA 25. The strong correlation results of the three UOAs mentioned previously, have been

verified by using a different dataset and a different methodology for Journal Ranking. By using

the SCImago SJR score [Scimago Journal and Country Rank, 2017] the strong correlations were

replicated, with the cases of UOA 5 and 18 returning even stronger correlation results. This can

hardly be a coincidence, and is a strong indication that both Research Income and the Journal

Rankings were good indicators in predicting the output for the REF 2014 process in the UOAs

that returned strong correlation results.

**Feature Work**

7.2

More on the Citations

To get a better understanding of the results it would be ideal if citation data is extracted from

multiple sources, such as Google Scholar and Scopus. This would provide a richer data set for

all UOAs and more reliable results.

7.3

More on the Journal Rankings

The SCImago Journal & Country Ranking List could be used to assess how the journal ranking

influences the REF results in more UOAs. By fully utilising this dataset together with a more

up-to-date ERA Journal Ranking List could provide more reliable results. Additionally given

more time it would be better for the analysis if all the journals in the REF datasets were

matched and given a score.

What would also be interesting to explore, would be to sample the REF datasets and form

training and validation datasets, and thus build a predictive model to whether journal rankings

could possibly be used in the next REF process as indicators. Ideally in this case data from

previous processes, such as the RAE could be included.

**Conclusion**

Analysis was performed on several features, in an attempt to identify whether they have any

effect on the output score produced by the REF.

Allocating the institutions to the Russell group and the non-Russel group, indicated that

there is a large variance between the REF outputs between the two. The Russell group uni-

versities had a better average of 4\* percentage papers produced. This can either suggest that

there is institutional bias in the REF process or that simply the Russell Group universities are

better.

When citations are considered, the results showed that there were variances in the correlation

results between UOAs, with some UOAs having stronger correlation than others. This can

suggest that there was a difference in the extent of how the citations were used as an indicator

of the academic performance of work in the assessment process.

The strong correlation results in several UOAs between the average income per submission

and the REF output, can provide an indication that higher research income institutions produce

higher quality work, or that possibly the assessment process was biased. Biased in the sense

that work submitted by higher research income institutions was assessed more leniently than

the other institutions.

For the journal ranking several UOAs returned very strong correlation results, with UOA 18

having the strongest correlation of R-value 0.95 with the SCImago Journal Ranking list. What

the results can imply in these UOAs is that the higher the quality of the paper the higher the

journal it is submitted to. Alternatively there is the possibility that the sub-panels responsible

for the strongy correlated UOA’s were biased, and papers published in higher ranked journals

were more leniently assessed in comparison to the papers published in the lower ranked journals.