Commonwealth Scholarship Commission in the UK

Successes and complexities: the outcomes of UK Commonwealth Scholarships 1960-2012

Annex one: Methodological notes

COMMONWEALTH SCHOLARSHIPS

This report was written by Dr Matt Mawer, Shireen Quraishi, and Rachel Day, and published in April 2016.

For further information regarding the CSC Evaluation and Monitoring Programme, please contact:

Commonwealth Scholarship Commission in the UK
Woburn House
20-24 Tavistock Square
London WC1H 9HF
UK
evaluation@cscuk.org.uk
http://bit.ly/cscuk-evaluation

Contents

List of figures

List of tables

Introductory note

Annex One: Methodological notes is an accompanying document to the following report:

Mawer, M., Quraishi, S. & Day, R. (2016). *Successes and complexities: the outcomes of UK Commonwealth Scholarships 1960-2012 - Full report*. London: Commonwealth Scholarship Commission in the UK

The annex provides additional commentary on several elements of the methodological and analytical approach taken with the full report. Additional questions about either survey or analytical details can be directed to the CSC Evaluation Team at the email address above.

Some terms are defined in the full report but are not repeated in this Annex; notably the overview of scholarship schemes. The reader is referred to the document cited above for fuller explanation.

For those wishing to cite the annex, we recommend the following:

Mawer, M., Quraishi, S. & Day, R. (2016). *Successes and complexities: the outcomes of UK Commonwealth Scholarships 1960-2012 – Annex one: Methodological notes*. London: Commonwealth Scholarship Commission in the UK

1. Regional composition

The CSC uses an in-house regional structure, dividing the areas of the Commonwealth into 8 regions. The core regional composition is constituted only by the 53 countries of the Commonwealth, but this typology has been extended to include those countries outside of the Commonwealth in which some respondents were resident. In some cases residency for a region would have been too small to be analytically useful (e.g. South America: N=1) and so the residency data for these respondents was excluded from most analysis (except measures of residency at home or abroad).

The final regional composition was:

1. **Australasia**
   Australia
   New Zealand

2. **Caribbean**
   Antigua and Barbuda
   Bahamas
   Barbados
   Belize
   Bermuda
   Dominica
   Falkland Islands
   Grenada
   Guyana
   Jamaica
   Montserrat
   Saint Kitts and Nevis
   Saint Lucia
   Saint Vincent and the Grenadines
   Trinidad and Tobago
   Virgin Islands (US)

3. **Europe**
   Belgium
   Cyprus
   Czech Republic
   Denmark
   Finland
   France
   Germany
   Gibraltar
   Greece
   Italy
   Lithuania
   Malta
   Netherlands
   Romania
   Sweden
   Switzerland
   United Kingdom

4. **North America**
   Canada
   United States

5. **Pacific**
   Fiji
   Palau
   Papua New Guinea
   Solomon Islands
   Tonga

6. **South Asia**
   Bangladesh
   India
   Maldives
   Pakistan
   Sri Lanka

7. **Southeast Asia**
   Brunei Darussalam
   China
   Hong Kong
   Japan
   Malaysia
   Philippines
   Republic of Korea
   Singapore
   Taiwan (Republic of China)
   Thailand

8. **Sub-Saharan Africa**
   Botswana
   Cameroon
   Ethiopia
   Ghana
   Kenya
   Lesotho
   Liberia
   Malawi
   Mauritius
   Mozambique
   Namibia
   Nigeria
   Rwanda
   Seychelles
   Sierra Leone
   Somalia
   South Africa
   South Sudan
   Swaziland
   Tanzania
   The Gambia
   Togo
   Tunisia
   Uganda
   Zambia
   Zimbabwe

2. Survey representativeness

An analysis of representativeness was conducted to explore the extent to which respondents to the CSC evaluation survey reflected Commonwealth Scholars and Fellows generally. It is important to note that this survey exercise was not a sample survey, but rather a census approach, and as such we did not set out to elicit responses from a representative subset of Commonwealth Scholars and Fellows. Nonetheless, it is still important to consider representativeness and any bias that may shape analytic results due to over- or under-participation by particular groups.

Through a comparison of certain variables in the survey respondents and general population of Commonwealth Scholars and Fellows we have generated a short analysis of representativeness.

A *very high* degree of representativeness between respondents and population was found in:

- **Selection scores and grades:** Selection process variables are very similar between the survey respondents and the general Scholar population, where such data is available (post-2004). For instance, there has not been disproportionate survey response from high-scoring candidates, or candidates with a strong initial development-potential profile.

- **Age at award uptake:** The survey respondents show very similar central tendency and dispersion for age at the uptake of their Commonwealth Scholarship.

- **Gender:** The proportion of male and female respondents in the survey sample is very similar to the population

The areas in which representativeness was *moderate-to-high* were:

- **Doctoral submission:** Doctoral scholarship holders that responded to the survey had submitted their thesis 'on time'[1] somewhat more frequently (+10%) than the population generally. Similarly, survey respondents had slightly lower average time to thesis submission than the population generally, by about 2 months.

- **Scholarship scheme:** Generally, Agency: Developed Scholars are over-represented in the survey respondents on a decade-by-decade basis. Fellows are somewhat over-represented in recent decades (1990s, 2000s) but under-represented in prior decades. Distance Learning Scholars are considerably under-represented: in the 2000s they constitute 10% less of the survey sample than they do the population of Commonwealth Scholars (8% vs 18% respectively).

- **Degree type**: Doctoral scholarships are over-represented in the survey respondents. Fellows are somewhat over-represented in recent decades, but under-represented in prior decades. There is a slight under-representation of postgraduate degrees in the 2000s, likely due to the under-representation of Distance Learning Scholars, who exclusively study postgraduate degrees.

The areas in which representativeness was *moderate* were:

- **Region of citizenship:** Two notable trends regarding region of citizenship are observable. Firstly, Australasian and North American (Canadian) Scholars are over-represented in the survey respondents across each decade of awards. Secondly, Scholars from South Asia and Sub-Saharan Africa are under-represented in the survey respondents, particularly in earlier decades of award holders (1960s, 1970s). In more recent decades the under-representation is very minor, whereas in the 1960s, for instance, the proportion of South Asian survey respondents was only 8%, whereas 33% of all award holders in that decade were from South Asia.

There were no areas in which the representativeness of the survey respondents was very poor: for instance, where a population group was entirely absent from the survey respondents or was very substantially over or under-represented (e.g. +/-30%). It should be noted that in some areas there are gaps in the data, particularly concerning historical award holders (esp. 1960s).

Two useful cautionary points can be drawn from the examination of representativeness. Firstly, the outcomes of Scholars and Fellows from the developing Commonwealth in the earlier decades of the

[1] For doctorates, the benchmark for submission is within 1 year of the conclusion of CSC full funding. Usually this equates to within four years, but suspensions or deferments can modify the exact duration. 'On time' would thus mean having submitted a thesis for examination within this period.

programme are likely to less evident in the survey data. As such, we are less able to explore the experiences and long-term impacts of early Scholars from the developing Commonwealth than for Scholars from, for instance, Canada, New Zealand, and Australia. The survey data can also misleadingly give the impression that in earlier decades the UK's Commonwealth Scholarships were dominated by Canada, New Zealand and Australia when, although there were proportionally more award holders from these countries than in recent decades, the majority of Scholarships and Fellowships in *all decades* have been taken up by recipients from South Asia and Sub-Saharan Africa. At their peak, North America and Australasia combined account for less than a quarter of Commonwealth Scholarships held within the decade, and only around 16% of *all Commonwealth Scholarships* between 1960 and 2012.

Secondly, within the survey data there is a small but notable over-representation of doctoral Scholars and related under-representation of those that have undertaken Master's courses; particularly by distance learning. The over-representation of the developed Commonwealth and doctoral Scholars is linked since the majority of awards to the high income, developed Commonwealth states have been for doctoral study, particularly in the early decades of programme. In consequence, however, when the survey data is analysed in aggregate, and not delineated by degree type, the outcomes of doctoral scholars are slightly over-weighted. An example is that employment outcomes may disproportionately reflect academic trajectories which are pursued by doctoral Scholars in far greater proportion than Scholars that undertook other degree types.

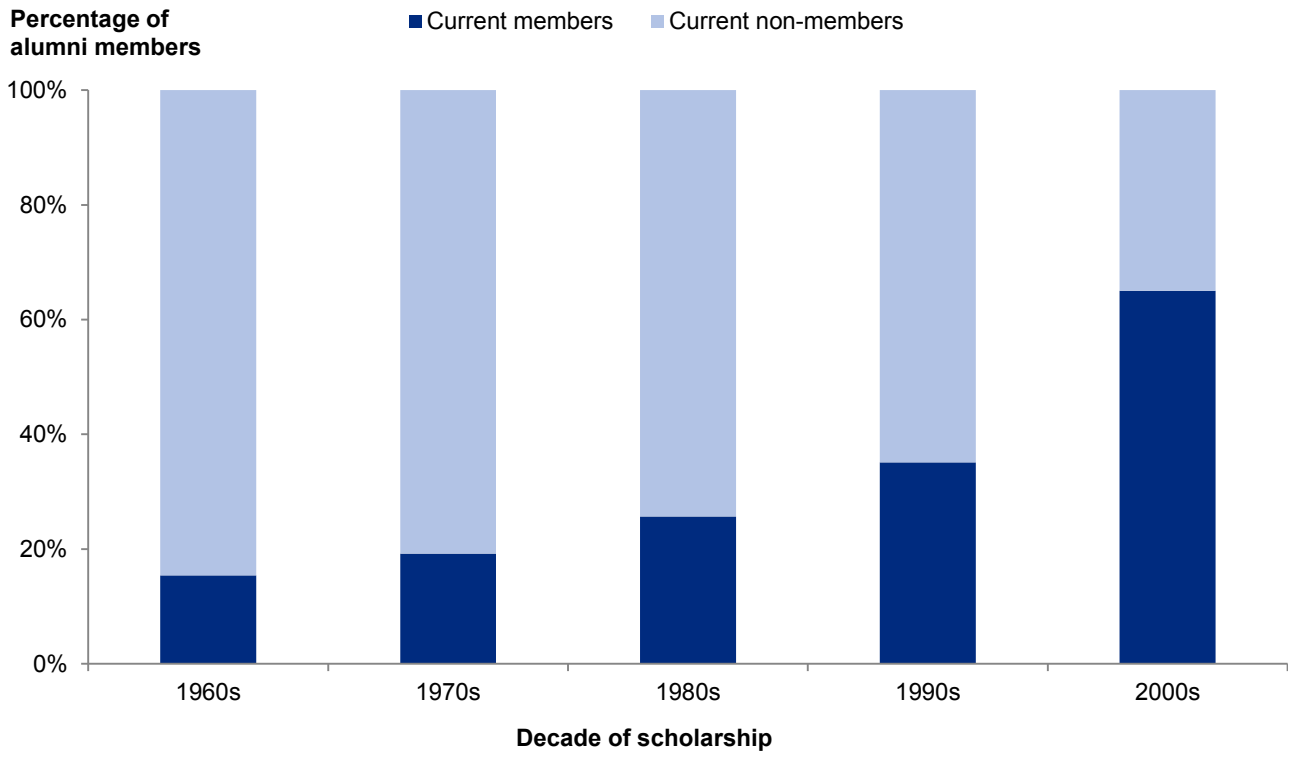
2.1. Survey non-response

Another dimension of representativeness is the influence of non-response bias on the survey findings.

The survey conducted as part of this research invited participation from members of the CSC alumni network. Although membership is open to all of the UK's Commonwealth Scholars and Fellows, not all take up this opportunity (for various reasons). Thus whilst the survey was conducted with an equal opportunity for all those within the *alumni network* to participate, not all those within the *population* of Commonwealth Scholars and Fellows had opportunity to participate if they were not members of the alumni network.

Alumni network membership has increased greatly in recent years due to a sustained drive by the CSC to trace and maintain contact with its alumni. As a result, current alumni members are more likely to have held Commonwealth awards in recent years, as Figure 1 illustrates.

Figure 1 **Membership of CSC alumni network by decade** (as of end 2015)



An effect of the differing alumni network membership is that the survey was sent to a greater proportion of Scholars and Fellows from some decades. Of Scholars from the 1960s, for instance, only 15% were eligible to receive a survey, whereas this figure was 65% for Scholars from the 2000s. These discrepancies partly

explain the lower absolute total of responses from earlier decades: e.g. only 6.4% survey respondents are Scholars from the 1970s, whereas 47.5% held a Scholarship in the 2000s.

Viewed from a different perspective, the survey data includes a lesser proportion of the overall population of Commonwealth Scholars and Fellows *within the decade* for the earlier decades of the programme.

Table 1 Survey respondents in comparison to all Commonwealth Scholars and Fellows

Decade of scholarship	Respondents	All Scholars and Fellows	Proportion that responded
1960s	75	2383	3.2%
1970s	133	3249	4.1%
1980s	215	4006	5.4%
1990s	354	4990	7.1%
2000s	993	6320	15.7%

Whether the differential alumni network membership across decades constitutes a problem of coverage bias depends on the characteristics of alumni network members in contrast to the population generally. If it is assumed that there is a relevant systematic difference between the alumni network members and the Scholar population generally (e.g. the former are the most motivated to translate scholarship gains into impact) then the surveys would suffer from coverage bias by excluding non-members. Conversely, if the alumni network members are broadly representative of the population on all measures of interest within the analysis then there would not be a coverage bias. At present there is no evidence available to the CSC to indicate any such systematic difference. Although it is possible to present scenarios showing bias – as in parenthesis above, for example – no comparative evidence is currently available to establish their plausibility. Additionally, it is likely that if such bias did exist, it is lower in recent decades and continues to lessen as more Scholars and Fellows join the alumni network and the disparity between the membership of the alumni network membership and the Scholar population decreases.

A second facet of non-response is the potential for differences between those who receive the survey in their propensity to respond. Groves et al. have noted that '…the key to nonresponse error is the correlation of the survey variables with response propensities' (2008: p2): in essence, the extent to which an outcome being measured through the survey is actually linked to whether a member of the sample is likely to respond. If this were the case then we might expect to find the survey results skewed toward a particular group that were more likely to respond. For instance, if high levels of social and economic impact were linked to a greater propensity toward completing the evaluation survey we might expect to find our survey results show relatively higher impact than is 'true' of the population.

As with the potential discrepancies between alumni network members and the general Scholar population, the CSC does not currently have research data supporting any scenario that shows plausible non-response bias. The levels of both reported catalytic impact and perceived gains are high, but this is a predicted outcome of an international scholarship programme, and there remain some 'low impact' cases (e.g. no reported social or economic impact) in the survey results. This does not, of course, mean no such bias exists, but rather than no evidence is available to support speculation.

If non-response bias was identified within the results it would remain somewhat difficult to make adequate adjustments. Statistical techniques for offsetting non-response are widely available (e.g. Blom, Lynn, & Jäckle, 2008: Blair, Czjaka, & Blair, 2014) but to a large extent rely on weighting by ancillary variables known to be linked to the primary variables under measurement. Since the only variables known for both population and survey respondents show a relatively robust level of representativeness in the survey it would not be particularly advantageous to use them for weighting. One exception could be region of citizenship, in which data could be weighted to reflect the under-representation of South Asian and Sub-Saharan African Scholars. In all cases, however, the problem of non-response bias is most significant when treating the survey data as a cohesive whole, which is often not the purpose of the analyses presented in the accompanying full report. Data are frequently disaggregated by region of citizenship, gender, degree type and so forth, as the differences in outcomes between subgroups are often as important as the outcomes across the Commonwealth Scholarship and Fellowship Plan as a whole.

3. Rescaling procedure

Due to minor changes between survey iterations, three questions needed rescaling to bring all data into the same scale format for subsequent collation. Rescaling and collating data was preferable for both simplicity and flexibility than developing new derived variables that could account for scale differences (such as a geometric mean).

Several rescaling options are available, of which those reported by Dawes (2002, 2008) and Preston and Colman (2000) are both straightforward and generally applicable. The two methods are mathematical interpolations into a 10-point scale from 5-point scale data. The Preston and Colman (2000) rescaling technique involves the following formula:

$$\frac{(\text{Rating} - 1)}{(\text{Scale points} - 1)} * 100$$

This yields a rating out of 100 based on the original rating and the number of scale points. The formula can also be adapted to give a rating out of 10 by replacing the final multiplication by 100 with a multiplication by 10. The Dawes (2002, 2008) rescaling technique involves interpolating the scale points as Table 2 indicates.

Table 2 Dawes procedure for interpolating 5-point into 10-point scales

5-point scale	10-point scale
1	1
2	3.25
3	5.5
4	7.75
5	10

The following procedure was followed to test the fit of both rescaling techniques:

1. Both Dawes rescaling technique and Preston and Colman rescaling technique were performed as necessary on items within data from the 2012 and 2013 survey iterations, leaving the 2014 and 2015 (10-point scale) data unaltered.

2. Descriptive statistics for the central tendency, dispersion, and distribution shape for all data (both rescaled and original) were taken. In the case of the original data, these statistics were calculated for each scale group (e.g. all the 5-point scales together, the 10 point scale separately) to avoid distortion.

3. Fit across the first four data moments - mean, variance, skewness, and kurtosis – was assessed

4. The rescaling method that produced the closest fit to the original data through each of the four moments was selected as the most appropriate for use in the final analysis

The Dawes (2002, 2008) procedure achieved the best fit and was been selected as the appropriate method. A more mathematically complex solution is detailed by Holmes and Mergen (2014), but the basic transformation suggested by Dawes fit our data well throughout and so we did not pursue this alternative.

4. Free-text analysis method

In addition to quantitative data, the survey also collected a substantial corpus of free text through questions answered in the respondents' own words. In particular, two questions were asked that were designed to generate concrete examples of development activities in which respondents had been involved: one question regarding socioeconomic impact and one regarding government policy influence. The structure of the questions was as follows:

> *'Thinking about the development areas you have chosen, please give a specific example of your influence on government policy'*

Some limited guidance was given on how these questions might be answered:

> *'For your example please give details of the following:*
>
> 1. *The activity (e.g. project run, book published, government policy implemented)*
> 2. *The outcome (e.g. decreased food poverty in x area, etc.)*
> 3. *The evidence (e.g. 25 individuals earning 15% more, literacy rates increased by 10% etc.)*
>
> *If you are not able to tell us the exact outcome or evidence, please give details about the intended outcomes or evidence.'*

To analyse the data, a free-text analysis strategy was developed based on coding responses and combining semantically similar codes to create a typology of activities undertaken by Commonwealth Scholars and Fellows.

Coding was conducted on the basis of identifying the type of the activities being described by respondents. We focused on what those activities involved, rather than what the outcomes of activities were. To 'code' in this case meant to read closely the text of the response and to summarise in a few words what was being described. As the question refers to outcomes and evidence, in addition to activities, it might theoretically be possible for further analysis to create a typology of outcomes, but our experience in this analysis would suggest that the data gathered through the current questions are not detailed enough to produce a credible interpretation of outcomes.

For each question the researchers independently coded several hundred responses to begin the analysis. When this coding was completed the researchers worked first independently and then together to consolidate the disparate codes into a series of categories that shared a common semantic base. For instance, codes relating to training, teaching, mentoring, guidance, etc. might all be grouped into a category called 'teaching and training'. Similarly, codes relating to starting a new business, setting up an academic research institute, and so forth, might be grouped into 'creating organisations'. The categories generated were then taken back to the data and another several hundred responses coded independently before again consolidating jointly. As more coding was conducted, the categories were revised, expanded, contracted, added or removed as appropriate. It was not intended that activities be coded into only one category: particularly in the case of project work, respondents' typically reported a combination of activities and so we did not attempt to force these into a single category.

The aim of categorisation was to generate a typology of categories with sufficient 'fit' to cover the vast majority of activities described, but whilst preserving the differences between 'types' of activity. The final typology is the result of several rounds of coding, categorisation and revision across both questions. It was decided at the end of the process to create one typology that covered both questions - rather than one per question – as there was significant overlap in activities. Additionally, there was a margin of error in which responses (particularly to the government policy question) were not being coded because they referred to activities that had no immediate bearing on government policy, but that could be coded under socioeconomic activity. The categories described in the Section 3.3 of the full report represent the major themes in survey responses.

5. Example of statistical procedures

Extensive statistical analysis of quantitative data collected by the survey has been conducted: this analysis underpins much of the interpretation and discussion in the full report. Rather than include a lengthy record of statistical procedures in this Annex we have documented (below) an exemplar of the approach we took where appropriate data was available.

5.1. Building a regression model for residency

To assess whether any of the variables on which we had collected either evaluation or administrative data could help to explain patterns in residency at home or abroad we sought to create a regression model. The model would help to explain which variables could be considered associated with residency outcomes such that a change in those variables would be associated with differing likelihood of residency at home or abroad.

As migration and 'brain drain' are more acute for lower income countries (Collier, 2015), the focus of our analysis specifically concerned lower income Commonwealth countries. As such, the basic parameters applied to this element of the data analysis were:

1. Respondent country of origin was not within the high income regions (North America, Australasia, or Europe)

2. Respondent had studied full-time in the UK on an academic scholarship; Fellowships, Split-site Doctorates,[2] and Distance Learners were excluded

3. Respondent was not funded under the Agency: Developed nominating route

The approach employed was logistic regression (due to the binary outcome): the outcome of interest was having reported a country of residence at the time of the survey either in the same region as the Scholar's country of citizenship ('home') or in a different region ('abroad'). We used regional aggregations rather than countries because many countries had only small numbers of survey respondents and would have been unsuitable for analysis.

Based on the data available to us, the following variables were then considered as 'predictors' of the residency outcome: 1) Gender, 2) region of citizenship, 3) degree type, 4) scholarship scheme, 5) candidate committee selection score (where available), 6) rating of employer supportiveness, 7) Counterfactual (UK study) rating, 8) Counterfactual (Another country study) rating, 9) current employment sector. In the terminology of the method, these are the 'predictor variables'. For the purposes of exploring the data, we hypothesised that one or more of these variables would be associated with a change in the likelihood of residing abroad.

Demographic variables and scholarship administration variables (1-5) were included in the analysis on the basis that these factors could be more easily influenced by CSC policy and thus any potential link with residency was of additional interest. Pre-scholarship variables (6-8) were included to explore whether current residency might be partially associated with either initial supportiveness of employers (e.g. by offering sabbatical or study leave) or respondent perception of their capacity to have studied abroad without a Commonwealth award. Finally, current employment sector (9) was included to examine whether any particular sector of employment was more closely associated with residence abroad.

5.1.1. Initial variable screening

We conducted univariate analyses of each possible predictor variable's association with residency outcome to aid model selection (following Hosmer & Lemeshow, 1989). By conducting this pre-screening it can be possible to remove some variables that appear to have no relation to outcomes and would thus be 'dead weight' in the analysis. In screening variables for model building Hosmer and Lemeshow (1989) recommend setting a low threshold for inclusion to avoid omitting potentially important variables or combinations of variables, usually α=0.25.

[2] While Commonwealth Split-site Scholarships involve full-time study in the UK, they are undertaken as part of doctoral study primarily based elsewhere and do not result in a qualification directly; Scholars return to undertake further study at their home institution after leaving the UK.

For the categorical variables, we conducted contingency table analysis and included a likelihood ratio chi square test of association.

Table 3: Univariate analysis of categorical residency variables

Variable	G^2	DF	Sig. (p)	Cramer's V
Gender	0.65	1	0.421	<0.001
Scholarship scheme	28.55	2	<0.001	0.030
Degree type	33.34	1	<0.001	0.040
Region of origin	4.99	3	0.173	0.005
Current employment sector	19.17	4	0.001	0.020

Using the low threshold for inclusion, all variables except gender were suitable for inclusion in the regression model. However, examining the Cramer's V effect size for each variable it is evident that none of the variables were closely associated with residency. Cramer's V is scaled between 0 and 1, but the highest effect size in table 1 is only 0.04, far below the standard used even for acknowledging a weak effect. Indications from the univariate analysis thus suggested that whilst there were several statistically relevant patterns in the data, none of these were likely to be analytically meaningful.

Similar results were found in examination of the continuous variables using univariate logistic regression, shown in Table 4.

Table 4: Univariate analysis of continuous residency variables

Model	G2	R2	Sig. (p)
Committee selection score	0.41	0.1%	0.524
Employer supportiveness	19.45	2.0%	<0.001
Counterfactual (UK)	3.23	0.3%	0.072
Counterfactual (Another country)	2.76	0.3%	0.097

Only committee selection score failed to qualify for inclusion based on the univariate logistic regression. However, like the categorical variables, there were strong indications in the screening that the continuous variables were likely to be of little analytical value in explaining residency. The largest R^2 was approximately 2%, with the remainder less than 0.5%, indicating that little variation in the residency outcome was associated with any of the continuous variables.

5.1.2. Model building diagnostics

Gender and committee score variables were excluded based on the univariate screening. Degree type and scholarship scheme were noted to be so closely related as to cause analytic problems within the regression. We concluded that exploring any variation between degree types was likely more useful than between scholarship schemes, of which several had already been excluded by our initial decision to focus only on developing country scholarships in the UK.

Diagnostic checks were performed to detect nonlinearity in the logit (through Box-Tidwell transformations), multicollinearity (through variance inflation factors), and empty or low cell count (see Menard, 2008). Region of citizenship was removed from the analysis at this stage as its inclusion yielded many instances of low cell counts: small group sizes for certain combinations of degree type, current employment sector, and region of citizenship would not have been suitable for analysis.

None of the Box-Tidwell transformations of continuous variables were statistically significant, indicating that there was no evidence of nonlinearity in the logit.

Table 5 Logistic regression of regional residency by Box-Tidwell (BT) transformations of continuous variables

Variable	Chi-Square	sig.(p)
Regression	62.62	<0.001
BT-Employer supportiveness	0.75	0.387
BT-Counterfactual (UK)	0.07	0.792
BT-Counterfactual (Another country)	1.13	0.288

Multicollinearity diagnostics were conducted by calculating the variance inflation factors for the variables included in the analysis. A useful benchmark is that variance inflation factors higher than five are a cause for concern: none of the variables reached this level of inflation.

Table 6 Variance inflation factors for variables included in the logistic regression model

Variable	Variance inflation
Employer supportiveness	1.11
Counterfactual (UK)	1.15
Counterfactual (Another country)	1.20
Current employment sector	
Academic	1.53
NGO	1.68
Other	2.05
Private	1.50
Public	-
Degree type	
Doctorate	1.22
Postgraduate	-

After the completion of model building diagnostics the remaining variables were: Degree Type, Current Employment Sector, Employer Support, Counterfactual (UK), and Counterfactual (Another country). The data available for analysis came from N=856 survey respondents.

5.1.3. Regression model

Logistic regression was conducted using deviance (-1, 0, +1) coding for categorical predictor variables and an increment of +/- 1 for continuous variables. The results showed several statistically significant coefficients for predictor variables: interpretation of these findings is detailed in Section 2.3.3 of the full report.

Table 7 Logistic regression of regional residency (outcome=Other) by employer supportiveness rating, counterfactual scores (UK and another country), current employment sector, and degree type

Variable	Coefficient	Standard error	95% Confidence interval	Sig. (p)
Constant	*-1.704*	*0.373*	*(-2.434, -0.973)*	*<0.001*
Employer Supportiveness	-0.058	0.037	(-0.131, 0.016)	0.128
Counterfactual (UK)	-0.015	0.049	(-0.111, 0.081)	0.761
Counterfactual (Another Country)	0.087	0.037	(0.015, 0.159)	0.017
Current employment sector				<0.001
Academic	0.520	0.182	(0.163, 0.876)	0.004
NGO	-0.140	0.291	(-0.709, 0.430)	0.631
Other	0.458	0.361	(-0.249, 1.166)	0.204
Private	-0.001	0.228	(-0.448, 0.446)	0.996
Public	-0.837	0.275	(-1.376, -0.298)	0.002
Degree Type				<0.001
Doctorate	-0.678	0.123	(-0.920, -0.437)	<0.001
Postgraduate	0.678	0.123	(0.437, 0.920)	<0.001

The overall model was statistically significant, with a final likelihood chi-square statistic of G=54.44 (DF=8, p<0.001). A total of 7.97% of variation in residency outcomes was associated with the predictor variables (the pseudo-R^2 statistic). Usually a regression analysis would expect to attain a higher R^2 if there was a strong association between predictor variables and outcomes, but our model was constructed in the context of exploratory data analysis with a series of variables that may either be directly associated or co-vary with related variables. An example of variables that might fit this profile are the Counterfactual (UK and Another Country), which logically cannot exert an influence on residency, but might co-vary with, for instance, socio-economic status or family networks abroad, which *may* be more directly associated with residency outcomes. The regression model developed does not show a *very strong* overall association (i.e. a high pseudo-R^2) between the predictor variables considered and residency outcome, but it does indicate several instances in which there is clear evidence for a link that is both statistically and analytically meaningful.

To further check the fit of the regression model we conducted several additional goodness-of-fit tests: the Pearson and Deviance chi square tests and the Hosmer-Lemeshow test (10 groups).

Table 8 Goodness-of-fit tests for logistic regression of regional residency

Test	DF	Chi-Square	P-Value
Deviance	670	628.42	0.873
Pearson	670	673.06	0.460
Hosmer-Lemeshow	8	13.49	0.096

Each of the goodness-of-fit tests reported a significant (p) value greater than the alpha threshold of 0.05, indicating that the hypothesis of adequate model fit should not be rejected.

All variables that were both potentially relevant and had sufficient data for analysis were included within our initial regression model and thus there was no further stage of model re-specification. If further information is available from the CSC's administrative and evaluation data collection in the future then the residency model may be expanded and the fit of various models, including alternative variables, could be developed.

References

Blom, A., Lynn, P., Jäckle, A. (2008). *Understanding cross-national differences in unit non-response: the role of contact data*. ISER working paper, Series no. 2008-01. Retrieved 6 March 2016 from: <http://hdl.handle.net/10419/92001>

Blair, J., Czaja, R., & Blair, E. (2014). *Designing surveys: A guide to decisions and procedures*. London: Sage

Collier, P. (2015). *Exodus: how migration is changing our world*. Oxford: Oxford University Press

Dawes, J. (2002). Five point vs. eleven point scales: does it make a difference to data characteristics. *Australasian Journal of Market Research*, 10(1), 39–47.

Dawes, J. (2008). Five point vs. eleven point scales: does it make a difference to data characteristics? *International Journal of Market Research*, 51(1), 61–77.

Groves, R., Brick, J.M., Couper, M., Kalsbeek, W., Harris-Kojetin, B., Kreuter, F., Pennell, B., Raghunathan, T., Schouten, B., Smith, T., Tourangeau, R., Bowers, A., Jans, M., Kennedy, C., Levenstein, R., Olson, K., Peytcheva, E., Ziniel, S. & Wagner, J. (2008). Issues Facing the Field: Alternative practical measures of representativeness of survey respondent pools. *Survey Practice*. 1 (3). Retrieved 4 March 2016 from <http://www.surveypractice.org/index.php/SurveyPractice/article/view/221/html>

Holmes, D. S., & Mergen, A. E. (2014). Converting survey results from four-point to five-point scale: a case study. *Total Quality Management & Business Excellence*, 25(1-2), 175–182. <http://doi.org/10.1080/14783363.2013.799330>

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley.

Menard, S. (2008). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks, Ca.: Sage

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15. <http://doi.org/10.1016/S0001-6918(99)00050-5>

Commonwealth Scholarship
Commission in the UK

Woburn House
20-24 Tavistock Square
London WC1H 9HF
UK
evaluation@cscuk.org.uk
http://bit.ly/cscuk-evaluation