Annex A

TECHNICAL ANNEX

ANNEX A
TECHNICAL ANNEX

Approach to index construction

This technical annex outlines how the indices of professionalism are constructed, and contains additional information on their distribution. The approach to scale construction in this study differs from that used by the Teaching and Learning International Survey (TALIS) for complex scales, such as their scales for job satisfaction and teacher efficacy, which weigh factors differently based on their contribution to an underlying latent variable. In this report, additive scales are created based on implementation of best practices, rather than complex scales based on latent variables. Following recommendations of the *TALIS 2013 Technical Report* (OECD, 2014b), similar scales are created using confirmatory factor analysis and test for overall fit and scalar invariance. In most cases, the scales exhibit a relatively good fit overall cross-nationally, but are not scalar invariant across all countries. We also find that additive component scales are very highly correlated to factor scales (~0.90+), while also having the added advantages of comparability. As such, we made the decision to work with the additive composite indices because they are more intuitive, comparable and have better distributions for subsequent analyses.

Index construction methodology

Following the theoretical literature, we outlined the best practices and policies that are shown to support teacher professionalism. Within each of the three domains (i.e. knowledge, autonomy and peer networks), we identified the TALIS questions that align to best practices and recoded them as binary variables. We conduct the index construction by ISCED level to decrease the possibility that variations in professionalism practices differ by school level. Additive composite indices were then created based on the literature, scaled to zero to five for comparability.

Following the *TALIS 2013 Technical Report* (OECD 2014b), we also tested how additive indices align to scales created by using structural equation modelling that treats each domain as a latent concept, rather than a total of observed practices. To carry out the structural equation models, we drew a random subset of 100 observations (teachers in the case of knowledge and peer networks and principals in the case of autonomy) from each country. This weighs each country equally in the construction of the scale. We then used structural equation modelling to predict latent factor scales for each domain, and tested the scale's goodness of fit overall and across all countries. As discussed in the literature review, we find that the two approaches are highly correlated and opt for the additive index approach.

Knowledge domain

Drawing on the literature, we focus on ten variables related to the types of knowledge teachers need to have to be successful, outlined in Chapter 1. We recode the knowledge variables to binary variables when appropriate such that a 1 represents higher requirements for professional knowledge and more support for professional learning. Although the literature suggests that a graduate-level degree is important to a teacher's knowledge base, we found that the overwhelming majority of respondents in TALIS reported that a Bachelor's degree is their highest degree (~90%), which left little variation in the index. Instead, we prioritise the variable for participation in a teacher education programme, along with other supports for professionalism.

Removing the question on highest level of education, the knowledge domain includes the following ten variables:

Table A.1 Knowledge domain variables

Sub-domain	Variables	Recoding
Pre-service education requirements	11. Did you complete a teacher education programme?	0 – No 1 – Yes
	12a) Content of the subject I teach included [in formal education or training].	0 – No / some subjects 1 – Yes, all subjects
	12b) Pedagogy of the subject I teach included [in formal education or training].	0 – No / some subjects 1 – Yes, all subjects
	12c) Classroom practice in the subject I teach included [in formal education or training].	0 – No / some subjects 1 – Yes, all subjects
Support for in-service professional learning (for teachers who participated in professional development in the last 12 months)	23) For the professional development in which you participated in the last 12 months, how much did you personally have to pay for?	0 – Some or all 1 – None
	24a) For the professional development in which you participated in the last 12 months, did you receive scheduled time for activities that took place during regular working hours at the school?	0 – No 1 – Yes
	24b) For the professional development in which you participated in the last 12 months, did you receive a salary supplement for activities outside regular working hours?	0 – No 1 – Yes
	24c) For the professional development in which you participated in the last 12 months, did you receive non-monetary support for activities outside working hours?	0 – No 1 – Yes
	25d) Considering the professionnal development activities you took part in during the last 12 months, to what extent have they included an extended time-period (several occasions spread out over several weeks or months)	0 – None 1 – Some, most or all
Support for practitioner research	21h) In the past 12 months, did you participate in individual or collaborative research on a professional topic of interest.	0– No 1 – Yes

Confirmatory factor analysis

To conduct the confirmatory factor analysis, in line with the methodology adopted by TALIS 2013, we first draw a random sample of 100 observations of teachers from each country, which weighs each country equally in the analysis. The Cronbach's alpha for all ten of the knowledge domain variables is lower than the acceptable threshold and the inter-item correlation is very low in all ISCED levels, which suggests that factor analysis may not be the best approach.

Table A.2 Cronbach's alpha of knowledge domain items, by ISCED level

ISCED level	Cronbach's alpha	Inter-item correlation
ISCED 1	0.494	0.0158
ISCED 2	0.490	0.016
ISCED 3	0.450	0.014

We then conduct a confirmatory factor analysis using structural equation commands to test the model fit of a scale developed from knowledge items. We examine three goodness of fit statistics – root mean squared error of approximation (RMSEA), comparative fit index (CFI) and standardised root mean squared residual (SRMR) – from the structural equation model, suggesting that a knowledge scale is a relatively good fit for a scale on the overall data.

Table A.3 SEM goodness of fit indicators of knowledge scales

Indicator	ISCED 1	ISCED 2	ISCED 3
RMSEA	0.050	0.065	0.069
CFI	0.932	0.904	0.872
SRMR	0.049	0.050	0.053

Additional analysis indicates that the factor loadings for the scale strongly emphasise formal education components (content, pedagogy and practice) over school-specific supports for ongoing professional development.

Additive index

We then create an additive index that weights all items equally and is scaled from zero to five. The additive scale has the advantage of being normally distributed and continuous.

Correlations between knowledge base scales

The scale is positively correlated with the other two scales; however, the correlation between the latent scales and additive indices is less than we would like – roughly 0.70. The relatively low correlation between the additive component and the latent variable approach comes from the differential weighting of factors in the construction of the scales. To disaggregate the pre-service and in-service professional development factors, we create two separate knowledge factors, one focusing on pre-service knowledge requirements and a second on in-school support for professional learning. The goodness of fit indicators suggest that this is generally a better fit for the data.

Table A.4 Goodness of fit indicators for a two-factor latent knowledge variable

Indicator	ISCED 1	ISCED 2	ISCED 3
RMSEA	0.029	0.043	0.055
CFI	0.978	0.959	0.921
SRMR	0.041	0.032	0.043

We then equally weigh each scale in the construction of a composite knowledge-base scale. This approach is more highly correlated with the additive composite scale, suggesting that a latent factor approach and additive composite index approach produce more similar indicators of professionalism when both pre-service and in-service professional learning are considered equally important to overall teacher professionalism. Because the theoretical literature on the topic consistently emphasises both pre- and in-service learning, we adopt the additive, composite approach that weighs both equally.

Table A.5 Correlations between knowledge domain scales and indices

	PCF	SEM – 1 latent	SEM – 2 latent
SEM – 1 latent variable	0.994		
SEM – 2 latent variables	0.943	0.937	
Additive composite	0.759	0.730	0.787

Autonomy domain

The variables on teachers' involvement in decision making are asked in only the principal questionnaire; to conduct the scale construction and analysis, we select a subsample of 100 principals from each country, as the questions in the autonomy scale are drawn from principal responses.

Table A.6 Autonomy domain variables

Domain	Question	Re-code
Autonomous decision making	Principal 18i) Do teachers have significant responsibility for choosing which learning materials are used?	0 – No 1 – Yes
	Principal 18j) Do teachers have significant responsibility for determining course content including (national/regional) curricula?	0 – No 1 – Yes
	Principal 18k) Do teachers have significant responsibility for deciding which courses are offered?	0 – No 1 – Yes
	Principal 18f) Do teachers have significant responsibility for establishing student disciplinary policies and procedures?	0 – No 1 – Yes
	Principal 18g) Do teachers have significant responsibility for establishing student assessment policies?	0 – No 1 – Yes

The Cronbach's alpha for these five items is quite high in all ISCED levels, roughly 0.75, which is above the conventional cut-off of 0.70.

Table A.7 Cronbach's alpha and inter-item correlation for autonomy

ISCED Level	Cronbach's alpha	Inter-item correlation
ISCED 1	0.79	0.09
ISCED 2	0.768	0.092
ISCED 3	0.797	0.102

Confirmatory factor analysis

Given the single factor score, we also conduct confirmatory factor analysis using structural equation modelling. The standardised coefficients are all very close to one another, ranging from roughly 0.50-0.70 – which suggests that, while some may be slightly more significant to autonomy, they may all also be equally weighted.

Table A.8 Goodness of fit indicators for autonomy scale, by ISCED level

Indicator	ISCED 1	ISCED 2	ISCED 3
RMSEA	0.127	0.167	0.151
CFI	0.946	0.887	0.923
SRMR	0.044	0.051	0.044

Additive composite index

Because the structural equation model suggests that all five components of the scale are likely to be equally weighted, there is support for constructing an additive index that takes each component or best practice as part of an additive, composite index.

Correlations between autonomy scales and additive index

Table A.9 Correlations between autonomy domain scales and indices

	PCF	SEM
SEM	0.980	
Component	0.999	0.976

The correlations are very high among all three approaches to measuring autonomy, which suggests that whether we use either a latent factor approach or a composite additive approach, the two are capturing the same phenomenon. We proceed with the creation of the additive scale due to its distributional benefits and ease of interpretation.

Peer networks

Based on the literature, the scale of peer networks draws on five practices: 1) induction; 2) mentoring; 3) peer feedback from direct observations; 4) existence of a personal professional development plan; and 5) participation in a professional learning community. The variables from TALIS are outlined below in Table A.10.

Table A.10 Variables in peer networks domain

Sub-domain	Question	Recoding
Induction	19a) I took part in an induction programme. (0/1)	0 – Did not take part in induction
		1- Took part in induction
Mentoring[1]	20a) I presently have an assigned mentor to support me. (0/1)	0 – Responded no to all mentoring questions
	20b) I serve as an assigned mentor for one or more teachers. (0/1)	
	21i) During the last 12 months, did you participate in mentoring and/or peer observation and coaching as part of a formal school arrangement? (0/1)	1 – Responded yes to at least one mentoring question
	31h) A mentor is appointed to help the teacher improve his/her teaching.	0 – Disagree or strongly disagree 1 – Agree or strongly agree
Peer feedback[2]	28a) In this school, the school principal, members of the school management team, assigned teachers or other teachers provide feedback based on direct observations of your teaching.	0 – No 1 – Yes
Development plan	31d) How strongly do you agree or disagree with the following statements: a development or training plan is established for teachers to improve their work as a teacher.	0 – Disagree or strongly disagree 1 – Agree or strongly agree
Network of teachers	21g) In the last 12 months, did you participate in a network of teachers formed specifically for the professional development of teachers?	0 – No 1 – Yes

Notes:
1. The variable takes a value of 1 if there is an affirmative response to any of questions 20a, 20b or 21i, or a response of "agree" or "strongly agree" to question 31h. It is set to missing if responses to all three of 20a, 20b and 21i are missing, as we place emphasis on the teacher's experience with mentoring at their school.
2. This variable is set to missing if there are missing values for all of TT2g28A2 through TT2g28A5.

The Cronbach's alpha on the five factors is relatively low, ranging from 0.44 to .054 and the inter-item covariance is 0.03-0.04.

Table A.11 Cronbach's alpha for peer networks domain, by ISCED level

ISCED level	Cronbach's alpha	Inter-item correlation
ISCED 1	0.438	0.030
ISCED 2	0.498	0.038
ISCED 3	0.539	0.045

Conducting a principal components analysis, we find that the various elements do reflect one underlying concept of strong professional networks. Given the single underlying factor suggested by the principal components analysis, we conduct confirmatory factor analysis using structural equation modelling on the five items suggested by the literature, including induction. Additionally, standardised coefficients are also quite close to one another (ranging from roughly 0.3-0.4) – which suggests that, while some may be slightly more significant to high peer networks overall, the differences are not substantial.

Table A.12 Goodness of fit indicators for the peer networks scale

Indicator	ISCED 1	ISCED 2	ISCED 3
RMSEA	0.044	0.019	0.049
CFI	0.965	0.992	0.965
SRMR	0.025	0.011	0.024

As shown in Table A.12, the goodness of fit statistics for the peer networks scale are quite good for the index at each ISCED level.

Additive component analysis

Because the structural equation model suggests that all five components of the scale are not substantively different from one another, there is also strong rationale for creating a composite index that weighs all items equally. We proceed with the creation of the additive scale due to its distributional benefits and ease of interpretation, creating a scale that naturally ranges from zero to five.

Correlations between peer networks scales

Table A.13 Correlations between peer networks domain scales and indices

Peer networks	PCF	SEM
SEM	0.99	
Additive composite	0.94	0.95

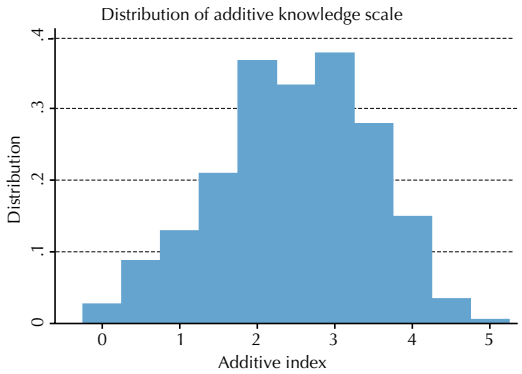
The correlations between the three scales are all quite high for all three scales, which suggests that whether we use either a latent factor approach or the additive composite approach, the two are capturing the same phenomenon.

Descriptive analysis of additive indices

This section provides an overview of the additive composite indices of teacher professionalism domains used in the analysis. From the histogram plots, it is clear that each of the domains has quite a different distribution. The descriptives also suggest more emphasis generally on knowledge base, followed by high peer networks, with the least emphasis on teachers' decision-making autonomy.

■ Figure A.1 ■

Distribution of knowledge base scale

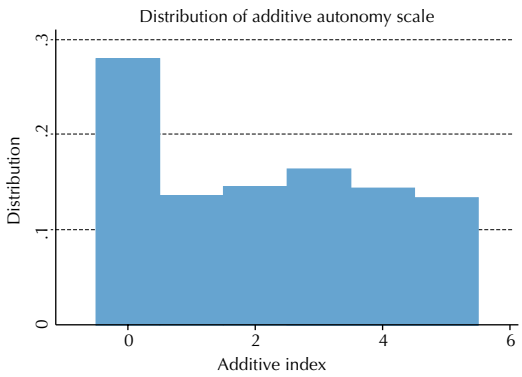


Source: OECD (2013), *Teaching and Learning International Survey (TALIS): 2013 complete database*, http://stats.oecd.org/index.aspx?datasetcode=talis_2013%20.

Table A.14 Distribution of knowledge base scale

ISCED level	Mean	Min	Max
ISCED 1	2.655	0	5
ISCED 2	2.657	0	5
ISCED 3	2.424	0	5

■ Figure A.2 ■
Distribution of autonomy scale

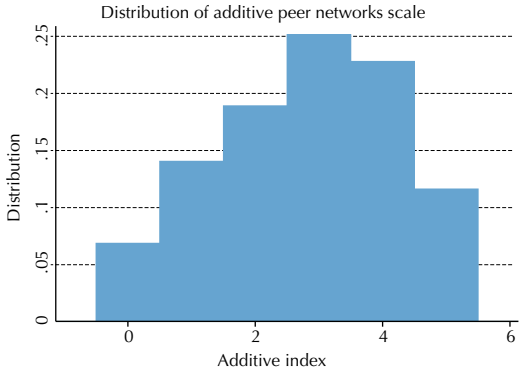


Source: OECD (2013), *Teaching and Learning International Survey (TALIS): 2013 complete database*, http://stats.oecd.org/index.aspx?datasetcode=talis_2013%20.

Table A.15 Distribution of autonomy domain scale

ISCED Level	Mean	Min	Max
ISCED 1	1.392	0	5
ISCED 2	1.978	0	5
ISCED 3	2.462	0	5

■ Figure A.3 ■
Distribution of peer networks scale



Source: OECD (2013), *Teaching and Learning International Survey (TALIS): 2013 complete database*, http://stats.oecd.org/index.aspx?datasetcode=talis_2013%20.

Table A.16 Distribution of peer networks domain scale

ISCED Level	Mean	Min	Max
ISCED 1	2.313	0	5
ISCED 2	2.772	0	5
ISCED 3	2.51	0	5

Treatment of missing values

Most of the variables in TALIS have some missing responses. Missing responses range from comprising 0% to up to 20% of all observations on some items. This is, in part, due to skip patterns in the TALIS questionnaire. Missing values need to be accounted for in scale construction wherever possible.

Observations with missing values are often not included in the construction of complex scales, which decreases the sample size and poses a risk of introducing selection bias, particularly if the missing pattern is not random. Missing values also pose a problem for additive scales, as a missing value mathematically does not contribute to the scale and, as a result, mathematically enters the scale equivalent to a response of zero, which is problematic for scale construction because we want to distinguish missing responses from negative responses.

To overcome the issue of missing data, we impute missing values with the school-mean for the questions where teachers are the respondents. This method is very effective at eliminating missing observations and has a minimal effect on overall distribution of responses to each question. The mean values hardly change, in the range of one-thousandth of a decimal point, while we are able to preserve many more observations. Understanding that mean imputation may underestimate variability in the values, which can lead to Type I error in interpretive analysis, we adjust the mean-imputed values by adding random variance equal to the amount of variability in the observed values prior to regression analysis.

With school-level variables, however, mean imputation was not possible, as schools were the primary sampling units. School-level variables were not imputed and missing values were list-wise deleted at the time of the analysis.

Regression analysis

Dependent variables

The four dependent variables were drawn from key items on TALIS. The unit of analysis for all dependent variables is the individual teacher.

Table A.17 Teacher professionalism outcome variables

Concept	Indicators
Status	I think that teaching is valued in society.
Satisfaction with work environment	I would recommend my school as a good place to work. I would like to change to another school if that were possible. I enjoy working at this school. All in all, I am satisfied with my job.
Satisfaction with profession	The advantages of being a teacher clearly outweigh the disadvantages. I regret that I decided to be a teacher. If I could decide again, I would still chose to work as a teacher. I wonder whether it would have been better to choose another profession.
Self-efficacy	To what extent do you believe that you can: Control disruptive behaviour in the classroom Make my expectations about student behaviour clear Get students to follow classroom rules Calm a student who is disruptive or noisy Craft good questions for my students Use a variety of assessment strategies Provide an alternative explanation or example when students are confused Implement alternative instructional strategies in my classroom Get students to believe they can do well in school work Help my students value learning Motivate students who show low interest in school work Help students think critically

For comparability, the dependent variables are all standardised, such that they have a mean of zero and a standard deviation.

Survey design and weights

The regression analyses in Chapter 3 use balanced repeated replicate survey weights, along with final teacher weights, using Stata 14 (StataCorp, 2015). The complex structure of the TALIS 2013 dataset necessitates specifying the survey characteristics of the dataset using the survey set command in Stata 14, which is done by using teacher weights, as well as the balanced repeated replicate weights (svy:). The primary sampling unit is the unique school identifier, with the brr option specified, using the 100 teacher replicate weights in TALIS (trwgt1-trwgt100). Fay's adjustment is set at 0.5. The svy prefix is used throughout the analysis, which ensures consistency in the application of final teacher weight and brr weights.

Regression models

The analyses of outcomes presented in Chapter 3 applies a two-level regression framework that accounts for the nested structure of the data, in which teachers are nested within schools, which are embedded within national education systems. For each of the dependent variables, the outcome is examined as a function of measures of teacher professionalism captured at the teacher level, with clustering at the school level. Across all model specifications, results showed that between-schools variation captures nearly all the variation between countries, making the clustering at the country level redundant.

The basic model predicts the outcome at the teacher level, as follows:

$$Outcome_{ij} = \beta_0 + \beta_1 P_{ij} + \beta_2 X_{ij} + \beta_3 W_j + e_{ij} + \varepsilon_j$$

Where, β_0 is the intercept for all teachers, P_{ij} is the relevant measure of teacher professionalism or its domain (measured at the teacher level for knowledge base and peer networks, but at the school level for autonomy), X_{ij} is a vector of teacher-level controls, and W_j is a vector of school-level controls. Some of the initial models also included a school mean on each of the professionalism measures, however, the final model places teacher professionalism at the teacher level to account for the individual-level variation. While in some model specifications school mean values for teacher professionalism were also tested as predictors of interest, results showed that including both teacher and school mean of teacher professionalism was not possible due to high multi-collinearity between these variables.

In addition to this basic model, the analysis also includes a number of control variables measured at the system level. This allows us to examine whether other system-level factors (i.e. male-female ratio, teacher pay, etc.) affect the relationship between teacher professionalism and outcomes. The set of models that tests these factors includes a vector of system-level covariates (Z_{jk}), including the male-to-female ratio, starting salaries and salary progression ladders, and contract type. These control variables are each included in the models individually, due to high correlations.

This extended model is structured as follows:

$$Outcome_{ij} = \beta_0 + \beta_1 P_{ij} + \beta_2 X_{ij} + \beta_3 W_j + \beta_4 Z_{jk} + e_{ij} + \varepsilon_j$$

The pooled multilevel model predicts each of the outcomes (status, job satisfaction, commitment and self-efficacy) as a function of individual and school variables, with system-level controls. As noted above, all analyses are done using the final teacher weight and brr weights.

Control variables

Our regression models control for important covariates. At the individual teacher level, controls include teacher gender and years of teaching experience. At the school level, controls include whether the school is public or private, the percentage of students who are socio-economically disadvantaged and an index of school climate, created by TALIS 2013, which captures how positive student-teacher relations are in the school.

Table A.18 Covariates in regression models

Level	Control variables	TALIS variable
Individual	Teacher gender	TT2G01
	Years teaching experience	TT2G05B
	Subject taught (coded as a series of four binary variables: 1) maths or science; 2) social sciences; 3) humanities or literature; 4) other)	TT2G15A-L
School	% of students in school from socio-economically disadvantaged homes	TC2G15C
	School climate – mutual respect (complex scale)	PSCMUTRS
	Public or private school	TC2G10
System	Relative salary compared to tertiary graduate in the labour force	OECD Education GPS (2014a)
	Test-based accountability system	Coded from Smith (forthcoming)

Table A.19 Regression models

Analysis type	Predictor variable	Regression model	Control variables
Overall	Teacher professionalism index	Pooled, two-level random intercepts model at the school level	Individual and school
Domain specific	Knowledge base scale	Pooled, two-level random intercepts model at the school level	Individual and school
	Autonomy scale	Pooled, two-level random intercepts model at the school level	Individual and school
	Peer networks scale	Pooled, two-level random intercepts model at the school level	Individual and school
Country-specific	Teacher professionalism index	Country-specific, two-level model with random intercepts at the school level	Individual and school
	Knowledge base scale	Country-specific, two-level model with random intercepts at the school level	Individual and school
	Autonomy scale	Country-specific, two-level model with random intercepts at the school level	Individual and school
	Peer networks scale	Country-specific, two-level model with random intercepts at the school level	
Alternate teacher professionalism checks	Teacher professionalism index	Pooled, three-level model with random intercepts for each school	Individual, school and additional teacher professionalism controls (salary and testing)

Robustness checks

A series of robustness checks were performed to ensure that the findings are robust to multiple model specifications and are not biased by the specifics of the sample, treatment of missing data or omitted variables.

To test whether the findings are biased by the cross-national sample, all cross-national models were also tested on random subsamples of 1 000 teachers drawn from each country. This ensured that all countries were equally represented in the cross-national study.

In addition to the TALIS data, we draw on system-level data to control for biases introduced by system-level factors. We draw on data from the 2014 Education GPS (OECD, 2014a), which is drawn from the OECD's annual *Education at a Glance* publication, and include indicators such as the male-female teacher ratio (an indicator of feminisation of the profession), various teacher salary measures, teacher-student ratios, teaching hours per year and the percentage of teachers by age bracket.

As shown in Table A.20, a series of system-level controls that may affect both teacher professionalism and outcomes of interest are also tested. Due to a high level of correlation between system-level variables, they are entered individually. The controls tested include feminisation of the profession, the salary ladder, economic development and the percent of teachers with a permanent contract. The results concerning the sign, significance and magnitude of the coefficients on teacher professionalism indices are robust to the inclusion of all the system-level controls tested.

Table A.20 Additional controls

Control variable	Source
Male-female teacher ratio	Education GPS (OECD, 2014a)
Salary ladder (ratio of salary at bottom to top of teacher pay scale)	Education GPS (OECD, 2014a)
Percent of teachers with permanent contract	Education GPS (OECD, 2014a)
GDP per capita	World Bank (2015)

Finally, to test whether the coding of missing values affects the findings, all models were run with controls for missing data and on smaller samples with no missing data.

References

OECD (2014a), *Education GPS*, http://gpseducation.oecd.org/

OECD (2014b), "TALIS 2013 Technical Report", OECD, Paris, http://www.oecd.org/edu/school/TALIS-technical-report-2013.pdf.

OECD (2013), *Teaching and Learning International Survey (TALIS): 2013 complete database,* http://stats.oecd.org/index.aspx?datasetcode=talis_2013%20.

Smith, W. (forthcoming), "Exploring educator based testing for accountability: national testing policies and student achievement", in T. Burns (ed.), *Modern Governance in Education: The Challenge of Complexity*, OECD Publishing, Paris.

StataCorp, (2015), "Stata Statistical Software: Release 14", StataCorp LP, College Station, TX.

World Bank (The) (2015), *World Development Indicators 2015*, http://data.worldbank.org/products/wdi.