# Table of Contents

1 Introduction .................................................................................................................. 6
   1.1 Background ................................................................................................................. 6
   1.2 Dominant Approaches to Understanding Education Quality ...................................... 8
       1.2.1 Human capital approach ....................................................................................... 8
       1.2.2 Human rights approach ....................................................................................... 9
       1.2.3 Criticisms of the human capital and human rights approaches .......................... 10
   1.3 Education Policy in South Africa Post-Apartheid ..................................................... 11
       1.3.1 Focus on Equity, Redress and Access ................................................................. 12
       1.3.2 Teacher Interventions and Curriculum Reform .................................................. 16
       1.3.3 Policy Lessons ..................................................................................................... 17
   1.4 Social Justice Approach to Education Quality ...................................................... 19
   1.5 Thesis Outline .......................................................................................................... 22

2 Tree of knowledge: A nonparametric approach to modelling primary school outcomes in South Africa .................................................................................................... 29
   2.1 Introduction ............................................................................................................ 29
   2.2 The Performance of Disadvantaged Schools in South Africa .................................. 33
   2.3 Empirical approach ............................................................................................... 35
       2.3.1 Regression trees .................................................................................................. 35
       2.3.2 Boosting and regularisation ................................................................................ 36
       2.3.3 Interpretation ...................................................................................................... 38
   2.4 Data .......................................................................................................................... 39
   2.5 Results ..................................................................................................................... 40
       2.5.1 From single to multiple tree regressions ............................................................. 40
       2.5.2 Tuning of model parameters ............................................................................... 41
       2.5.3 Relative influence of control variables and partial dependence plots ............... 43
   2.6 Identification of important interactions .................................................................... 48
   2.7 Robustness checks .................................................................................................. 53
       2.6.1 Sensitivity to dropping observations .................................................................. 53
       2.6.2 Comparisons across independent data sets ....................................................... 54
       2.6.3 Comparisons with alternative modelling approaches ...................................... 55
       2.6.4 Regression Tree Modelling with Clustered Data ................................................ 58
   2.8 Concluding Remarks ............................................................................................... 61
   Appendix to Chapter 2 .................................................................................................. 63

3 A question of efficiency: decomposing South African reading test scores using PIRLS 2006 ............................................................................................................. 67
   3.1 Introduction ............................................................................................................ 67
   3.2 Methodology .......................................................................................................... 69
       3.2.1 Oaxaca-Blinder decomposition ......................................................................... 69
       3.2.2 Semi-parametric decomposition ....................................................................... 72
   3.3 Data and summary statistics .................................................................................. 77
4 Balancing Act: A Semi-parametric approach for determining the local treatment effect of school type with an application to South Africa

4.1 Introduction
4.2 Motivation and Methodological Approach
4.2.1 Study Design: The effect of attending a former advantaged school
4.2.2 Potential Outcomes Framework
4.2.3 Inducing randomness: Creating comparable school and student groups
4.2.4 Post-matching estimation strategy
4.3 Data description
4.4 Empirical results
4.4.1 Matching students
4.4.2 Matching schools
4.4.3 Estimates of the treatment effect
4.4.4 Sensitivity analysis
4.5 Regression meets matching and propensity score weighting
5. Concluding remarks
Appendix for chapter 4

5 Learn to teach, teach to learn: A within-pupil across-subject approach to estimating the impact of teacher subject knowledge on South African grade 6 performance

5.1 Introduction
5.2 Policy context and previous findings in South Africa
5.3 Data and Descriptive Statistics
5.4 Estimation strategy: correlated random errors model
5.5 Results
5.5.1 Base results
5.5.2 Heterogeneous effects across student sub-groups
5.5.3 Returns to teacher and classroom characteristics
5.5.4 Correction for teacher unobservables
5.5.5 Fixed effects estimation
5.6 Conclusion
Appendix to Chapter 5

6 Compulsory tutorial programmes and performance in undergraduate microeconomics: A regression discontinuity design

6.1 Introduction
List of Tables

Table 1.1: Per-learner spending in public ordinary schools by province for selected years between 1998 and 2012 ................................................................. 13
Table 1.2: Actual provincial allocation per learner against national targets, 2012/13 (Rand) .......... 14
Table 1.3: Median departmental and total school NPNC spending per student (ZAR) ............. 15
Table 2.1: Predictive performance across model parameters ............................................ 43
Table 2.2: Summary of the relative contributions (%) of controls for boosted regression tree models of numeracy and literacy test scores ......................................................... 44
Table 2.3: Strongest two-way interactions for Grade 4 numeracy and literacy BRT models ........ 50
Table 2.4: Most influential predictors across BRT models of numeracy score using sub-samples of the NSES (2008) data ................................................................. 54
Table 2.5: Most influential variables in BRT models of numeracy across the NSES Grade 5 (2009) and SACMEQ Grade 6 (2007) datasets ................................................................. 56
Table 2.6: Model performance of competing approaches using training and test data splits .......... 57
Table 2.7: Variable importance across boosting and random forest models of numeracy and literacy ................................................................. 59

Table 3.1: Aggregate decomposition results ........................................................................ 82
Table 3.2: Detailed aggregate decomposition results ............................................................. 85
Table 3.3: Sensitivity checks based on propensity score selection rules, alternative model specifications and matching procedures ................................................................. 87
Table 3.4: Detailed decomposition results for different model specifications .................... 89

Table 4.1: National benchmarks for selected school and classroom resources .................... 109
Table 4.2: Estimated sample and population treatment effects of attending an EAT school ...... 122

Table 5.1: Cross-sectional regressions ................................................................................. 145
Table 5.2: Correlated random effects models ....................................................................... 146
Table 5.3: Correlated random effects models across sub-samples .................................... 148
Table 5.4: Correlated random effects models across different school sub-systems ............ 151
Table 5.5: Returns to other teacher and classroom characteristics ........................................ 155
Table 5.6: Correlated random error model results using the ST sample ............................ 157
Table 5.7: Student fixed effects estimation results ............................................................... 159

Table 6.1: Comparison of compulsory and non-compulsory tutorial attendance groups .......... 179
Table 6.2: Regression results for tutorial attendance and performance ............................. 182
Table 6.3: Non-parametric results ...................................................................................... 183
Table 6.4: Sensitivity checks ............................................................................................... 186
**Table of Figures**

Figure 1.1: A basic systems model of school effectiveness .................................................. 9
Figure 1.2: A simple context-led model for conceptualizing quality of education ................ 21

Figure 2.1: Kernel densities of grade 4 numeracy scores in 2008, by former school department... 41
Figure 2.2: Example tree structures and partial plots from a BRT model.................................. 42
Figure 2.3: Partial dependence plots for the nine most influential variables in the model for grade 4 numeracy ......................................................................................... 46
Figure 2.4: Partial dependence plots for the nine most influential variables in the model for grade 4 literacy ........................................................................................................... 47
Figure 2.5: Joint partial dependence plots illustrating two-way interactions from Grade 4 numeracy model ................................................................................................................ 52
Figure 2.6: RE-EM regression tree for Grade 4 numeracy ................................................... 60

Figure 3.1: Reading score distributions, by school type ....................................................... 80

Figure 4.1: Reading test score distribution by school test language ....................................... 114
Figure 4.2: Propensity score distribution across school test language ..................................... 116
Figure 4.3: Difference in standardised means of student and home background covariates, pre- and post-reweighting ....................................................................................... 118
Figure 4.4: Difference in standardised means of school, classroom and teacher covariates, pre- and post-reweighting ....................................................................................... 119
Figure 4.5: Difference in standardised means of school, classroom and teacher covariates not included in coarsened exact matching ............................................................ 121
Figure 4.6: School treatment effect (SATO) as a function of balance in school level covariates ... 124

Figure 5.1: Student performance by school SES quintile ...................................................... 149
Figure 5.2: Teacher performance by school SES quintile ...................................................... 149
Figure 5.3: Returns to teacher knowledge by performance quintile and school wealth group .... 152