Mokoro - Tanzania

# Data

lgacontrols: 89 variables ranging from population, poverty, goverment expenditure, data on educational system. Data ranges from 2015-2019. Not all data has complete series.

# Payments

payments\_alldrs: 9 variables. Payments in dollars per DLI between 2015 and 2019, disaggregated at council level (LGA).

Between 184 and 185 LGA are elegible, although not all of them receive funds. Not all DLI are paid across range of years.

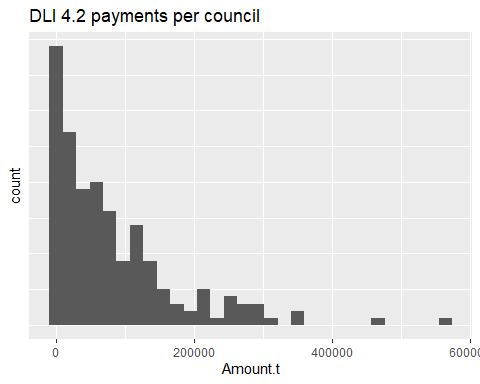
payments\_alldlrs.c: DLI payments are aggregated at council and DLI (losing time dimension) to use in the analysis of DLI 4 - PTR (only data on 2015 and 2019, a before/after).

About DLI 4.2: Average amount received: 82016$, std. dev: 92569; 33 councils received 0. Quartiles: 1st: 14782, 3rd: 115000

DLI 4.2 over time: many councils with 0; year 2018 is the highest avarage: 33880, std. dev 65192. years 2016 and 2017 average lower: 16141, 11207, respectevely.

DLI payments per year - potential councils

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2016 | 2017 | 2018 | 2019 |
| **2.2** | 184 | 0 | 0 | 0 |
| **2.3** | 0 | 0 | 185 | 185 |
| **3.2** | 184 | 184 | 184 | 184 |
| **4.2** | 184 | 184 | 184 | 185 |
| **7.1** | 0 | 184 | 185 | 185 |
| **7.2** | 0 | 0 | 185 | 185 |
| **8.2** | 0 | 0 | 0 | 185 |



Aggregated DLI 4.2 payments - Descriptive statistics

|  |  |  |  |
| --- | --- | --- | --- |
| count | Mean.Amount.t | SD.Amount.t | Councils without receiveing |
| 184 | 82016 | 92569 | 33 |

DLI 4.2 payments per year - Descriptive statistics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | count | Mean.Amount | SD.Amount | Councils without receiveing |
| 2016 | 184 | 16141 | 48358 | 131 |
| 2017 | 184 | 11207 | 29187 | 133 |
| 2018 | 184 | 33880 | 65192 | 82 |
| 2019 | 185 | 20676 | 47083 | 127 |

# DLR 4 - Pupil teacher ratio

Pupil teacher ratio ranges from 0 to 884 studentes per teacher. This suggests measurement errors. Setting outliers at <1.5% and >98.5%, excluded values are <8.8 and > 113.82.

PTR - Descriptive statistics

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | NA’s |
| 0 | 36.7 | 49.8 | 52.6 | 66.8 | 884 | 4 |

PTR is aggregated at LGA level because there is not match of 18172 of council/school names.

Analysing PTR over time (2015 and 2019), we observe median and mean are higher in the in 2019 roughly by 5.5 pupils.

PTR by year without outliers - Descriptive statistics

|  |  |  |  |
| --- | --- | --- | --- |
| year.ptr | Median.ptr.c | Mean.ptr.c | SD.ptr.c |
| 2015 | 50.6 | 49.4 | 10.6 |
| 2019 | 56.1 | 54 | 12 |

This is explained by an important raise in students population (15.5% on average) and a smaller increase in teacher population (4.7%).

Total number of students per council across years

* **2015**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 9756 | * 27763 | * 45079 | * 48421 | * 61432 | * 192743 |

* **2019**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 10524 | * 32053 | * 52079 | * 55807 | * 72915 | * 226878 |

Total number of teachers per council across years

* **2015**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 161 | * 617 | * 916 | * 1037 | * 1231 | * 5497 |

* **2019**:

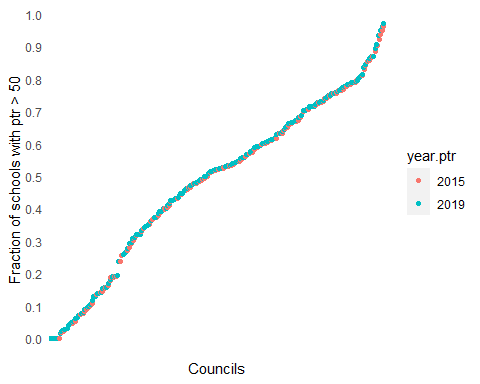
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 199 | * 623 | * 960 | * 1086 | * 1271 | * 5936 |

School PTR across councils have deteriorated between 2015 and 2019. The overall average of schools with a PTR over 50 (which was the standard in 2015) growths from 42.1% to 54.4% and dispersion, measured by the standard deviation, also increases form 23.4% to 26.1%.

The following plot shows the fraction of schools above the acceptable threshold for 2015 (>50 students) in both years, ranging from 0 to 97.25%.

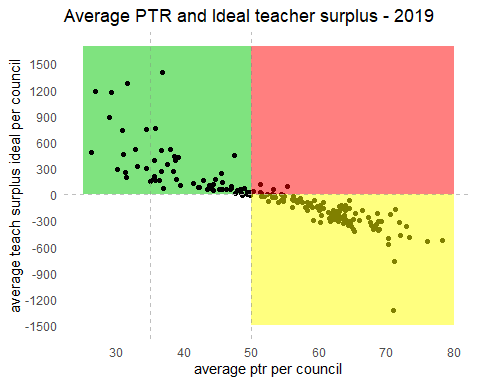
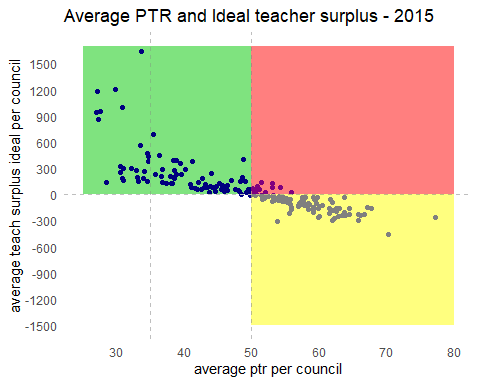
Share of schools with PTR > 50 - Descriptive statistics

|  |  |  |  |
| --- | --- | --- | --- |
| year.ptr | Median | Mean | SD |
| 2015 | 0.476 | 0.421 | 0.234 |
| 2019 | 0.576 | 0.544 | 0.261 |



An ideal surplus of teachers per council was computed as , which gives an idea of the ability of the council to achieve equaty in the distribtion of teachers. A major caveat of this calculation relies on the fact that students cannot follow the same ideal distribution due to household location, school resources and population age structure. However, this allows to understand the magnitude of gaps to be covered.

The following plots shows, for 2015 and 2019, the average PTR per council and the average ideal teacher surplus. The green area capture those councils under the PTR threshold, while the red are those which, having an ideal surplus, are above the PTR cutoff. The yellow area represents those councils that are both above PTR threshold and with teacher’s deficit.



There is a strong negative linear correlation (-.79 and -.86) across both years. Both tails, in the negative and positive side, show different patterns.

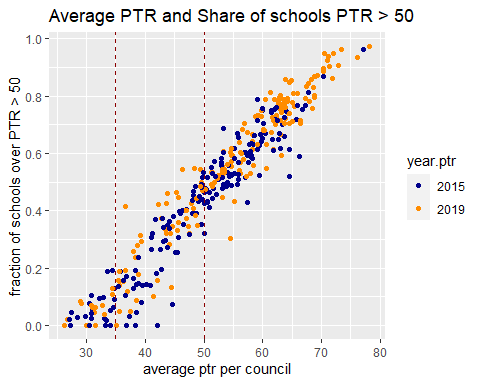
Correlation PTR and teacher ideal distribution by council

|  |  |
| --- | --- |
| year.ptr | cor |
| 2015 | -0.7943758 |
| 2019 | -0.8583754 |

There is a very strong correlation between average PTR and share of schools with PTR above 50 per council,as shown in the following graph: .95 and .97 in 2015 and 2017.

Correlation PTR and share of schools above PTR >50

|  |  |
| --- | --- |
| year.ptr | cor |
| 2015 | 0.9532881 |
| 2019 | 0.9662395 |



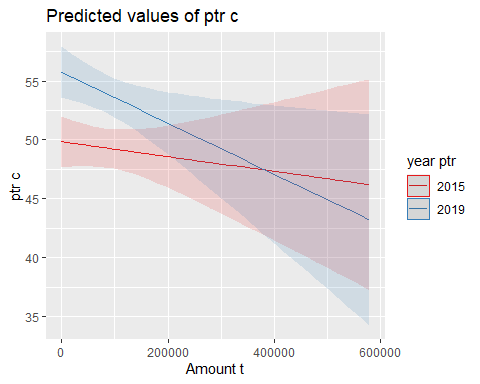
Turning to the analysis of PBR, we model a mixed-effects ANOVA to account for between- and within- council mean PTR differences. Onyl the year year and the interaction between year and DLI amount are significant.The relative magnitudes of the sums of squares indicates that the year term explains much more variation of PTR than the interaction term. Plotting the interaction predicted values we observe differences in terms of PTR across years when councils receive smaller values, which become insignificant later. To estimate differences, we use linear mixed-effects models, which are summarised below.

Analysis of Variance - Type 3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Sum Sq | Mean Sq | NumDF | DenDF | F value | Pr(>F) |
| **Amount.t** | 27.3 | 27.3 | 1 | 182 | 2.5 | 0.116 |
| **year.ptr** | 27545 | 13772 | 2 | 180 | 1257 | 1.1e-106 |
| **Amount.t:year.ptr** | 185 | 185 | 1 | 179 | 16.9 | 0.0000611 |

Model parameters

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | df | t value | Pr(>|t|) |
| **Amount.t** | -0.00000624 | 0.00000902 | 198 | -0.691 | 0.49 |
| **year.ptr2015** | 49.8 | 1.11 | 199 | 44.7 | 1.19e-105 |
| **year.ptr2019** | 55.7 | 1.11 | 198 | 50.1 | 2.07e-114 |
| **Amount.t:year.ptr2019** | -0.0000154 | 0.00000375 | 179 | -4.11 | 0.0000611 |



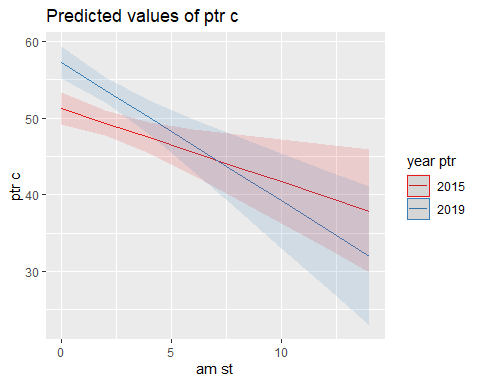
We also normalise councils’ received amounts by student, such as in and run the same ANOVA model. In this case, all terms, including the DLI amount, are significant, while the sum of squares suggest the year term is relatively higher in terms of explaining PTR. The amount/student estimate is significant and shows a negative sign, which suggests an association between amount/student and reducing PTR.

Analysis of Variance - Type 3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Sum Sq | Mean Sq | NumDF | DenDF | F value | Pr(>F) |
| **am.st** | 170 | 170 | 1 | 202 | 15.9 | 0.000092 |
| **year.ptr** | 31416 | 15708 | 2 | 184 | 1469 | 5.42e-114 |
| **am.st:year.ptr** | 307 | 307 | 1 | 199 | 28.7 | 0.000000236 |

Model parameters

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | df | t value | Pr(>|t|) |
| **am.st** | -0.957 | 0.336 | 220 | -2.85 | 0.00476 |
| **year.ptr2015** | 51.3 | 1.06 | 208 | 48.3 | 5.27e-115 |
| **year.ptr2019** | 57.3 | 1.06 | 207 | 54.2 | 2.36e-124 |
| **am.st:year.ptr2019** | -0.849 | 0.158 | 199 | -5.35 | 0.000000236 |



# DLR 6

## National examns

Correlation between # students and school score

|  |  |
| --- | --- |
| exam | cor |
| PSLE | 0.02908979 |
| SFNA | -0.1479596 |

Descriptive statistics - Ammount DLR

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | NA’s |
| 15527 | 44815 | 106906 | 134695 | 187638 | 851394 | 308 |

* **2016**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 16667 | * 36062 | * 36991 | * 52187 | * 41064 | * 376393 |

* **2017**:

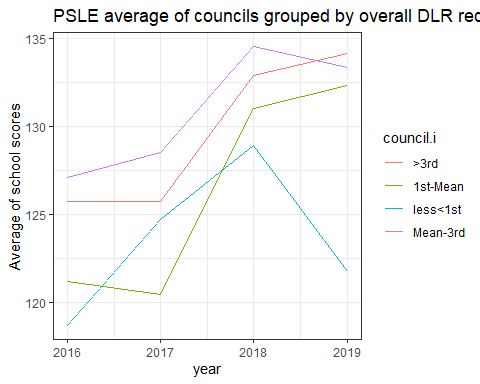
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 27206 | * 51048 | * 77206 | * 98195 | * 130236 | * 324206 |

* **2018**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 15527 | * 95789 | * 139508 | * 163823 | * 204551 | * 851394 |

* **2019**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Min. | * 1st Qu. | * Median | * Mean | * 3rd Qu. | * Max. |
| * 38168 | * 140865 | * 204447 | * 224573 | * 279948 | * 791257 |



Fixed effects models at council level - aggregating all money, considering learning outcomes (tests) as a impact.

## Twoways effects Within Model  
##   
## Call:  
## plm(formula = score ~ am, data = nat.concil, effect = "twoways",   
## model = "within", index = c("council", "year"))  
##   
## Unbalanced Panel: n = 182, T = 1-4, N = 696  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -30.017990 -3.091990 -0.023131 3.678247 26.410176   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## am 0.0000114244 0.0000039912 2.8624 0.004377 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 29642  
## Residual Sum of Squares: 29173  
## R-Squared: 0.015811  
## Adj. R-Squared: -0.3412  
## F-statistic: 8.19327 on 1 and 510 DF, p-value: 0.0043775

## Estimate Std. Error t-value Pr(>|t|)   
## 2016 119.25281 0.62732 190.10 < 2.2e-16 \*\*\*  
## 2017 122.32889 0.70491 173.54 < 2.2e-16 \*\*\*  
## 2018 130.82457 0.86389 151.44 < 2.2e-16 \*\*\*  
## 2019 131.37612 1.05566 124.45 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Random effects models at council level

## Twoways effects Random Effect Model   
## (Swamy-Arora's transformation)  
##   
## Call:  
## plm(formula = score ~ am, data = nat.concil, effect = "twoways",   
## model = "random", index = c("council", "year"))  
##   
## Unbalanced Panel: n = 182, T = 1-4, N = 696  
##   
## Effects:  
## var std.dev share  
## idiosyncratic 57.202 7.563 0.280  
## individual 142.250 11.927 0.696  
## time 4.849 2.202 0.024  
## theta:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.  
## id 0.4644675 0.6977630 0.6977630 0.6929966 0.6977630 0.6977630  
## time 0.7416987 0.7445720 0.7520049 0.7480923 0.7532879 0.7532879  
## total 0.4476793 0.6397208 0.6412462 0.6392554 0.6451080 0.6457621  
##   
## Residuals:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -41.563 -10.154 -1.265 0.245 9.722 50.145   
##   
## Coefficients:  
## Estimate Std. Error z-value Pr(>|z|)   
## (Intercept) 1.2558e+02 2.0246e-01 620.25 < 2.2e-16 \*\*\*  
## am 1.3903e-05 4.9868e-07 27.88 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 162230  
## Residual Sum of Squares: 155470  
## R-Squared: 0.067199  
## Adj. R-Squared: 0.065855  
## Chisq: 30.1907 on 1 DF, p-value: 3.9159e-08

## gamma nu   
## 0.9502553 0.9778201

Random effects models adding covariates (in lme4 - more flexible) - models don’t converge

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [  
## lmerModLmerTest]  
## Formula:   
## score ~ am + year + pop2019 + urbrur + povertyregion + dropout\_uw2015\_pct +   
## (1 + year | council)  
## Data: nat.concil  
##   
## REML criterion at convergence: 4401.6  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.7584 -0.4798 -0.0133 0.5248 3.3484   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr   
## council (Intercept) 64.66400169 8.041393   
## year 0.00003413 0.005842 -0.53  
## Residual 64.75501883 8.047050   
## Number of obs: 582, groups: council, 152  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) -9.501e+03 7.840e+02 4.588e+02 -12.120 < 2e-16 \*\*\*  
## am 8.550e-06 4.379e-06 5.023e+02 1.953 0.051427 .   
## year 4.768e+00 3.888e-01 4.588e+02 12.265 < 2e-16 \*\*\*  
## pop2019 1.443e-05 3.910e-06 1.472e+02 3.691 0.000313 \*\*\*  
## urbrurUrban 1.029e+01 2.168e+00 1.491e+02 4.749 0.00000476 \*\*\*  
## povertyregion 4.447e+00 1.290e+01 1.505e+02 0.345 0.730811   
## dropout\_uw2015\_pct -1.927e-01 4.999e-01 1.474e+02 -0.385 0.700436   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) am year pp2019 urbrrU pvrtyr  
## am 0.630   
## year -1.000 -0.630   
## pop2019 -0.033 -0.047 0.031   
## urbrurUrban 0.012 0.016 -0.012 -0.044   
## povertyregn 0.008 0.015 -0.011 0.211 -0.067   
## drpt\_w2015\_ -0.005 -0.015 0.005 -0.088 0.245 -0.534  
## fit warnings:  
## Some predictor variables are on very different scales: consider rescaling  
## convergence code: 0  
## unable to evaluate scaled gradient  
## Model failed to converge: degenerate Hessian with 2 negative eigenvalues