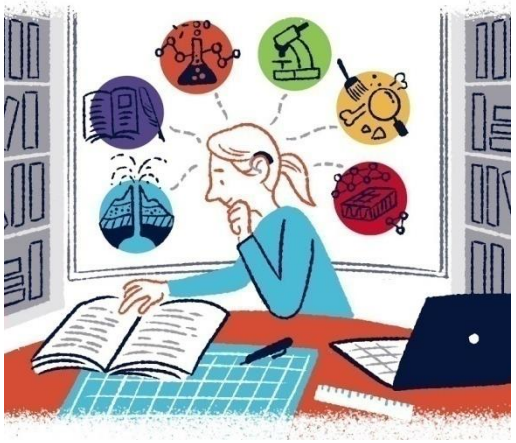
UvA



Inspectie van het Onderwijs
Ministerie van Onderwijs, Cultuur en Wetenschap

Predictability in Future Results of Primary Schools

Andreas, Zena and Thomas

University of Amsterdam, 21-10-2021

Outline

Are future changes in learning results predictable?

Situation

- The inspectorate of education provided data to look at the differences in learning results of students in primary school. Schools with a low proportion of students with an acceptable learning result could be low quality schools.
- The inspectorate of education wants to identify these low quality schools
- Important variables:
- The proportion 2F scores, representing the proportion of students with an acceptable learning result.
- Schoolweging, which is a demographic that contains all sorts of information about the situation that a child grows up in.

Complication

- Predict whether schools will become low quality schools in the future. Also how far in the future would be predictable and which indicators should be looked at.

Request

- The inspectorate of education is mainly interested in the predictability of changes in learning results over time. The most important questions are:
- Are future changes in learning results predictable?
- What indicators should they be looking at?
- How far in the future can the model predict?

Main Findings

Observed and Predicted difference on Test Data (year 2016 – 2017)

- Trained models on year 2015-2016
- 3 models:
  - 1. Polynomial model containing 3 variables
  - 2. Linear model containing 4 variables
  - 3. Polynomial model containing 12 variables

Figure 8: Observed and Predicted Difference in prop 2F Score for year 2016-2017 and 2017-2018 (Model with 4 Variables)

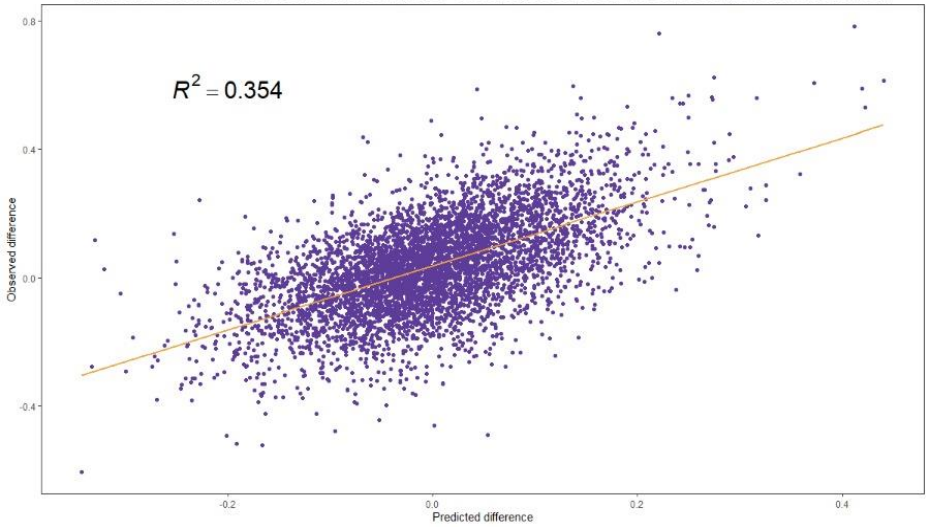


Figure 7: Observed and Predicted Difference in prop 2F Score for year 2016-2017 and 2017-2018 (Simple Model)

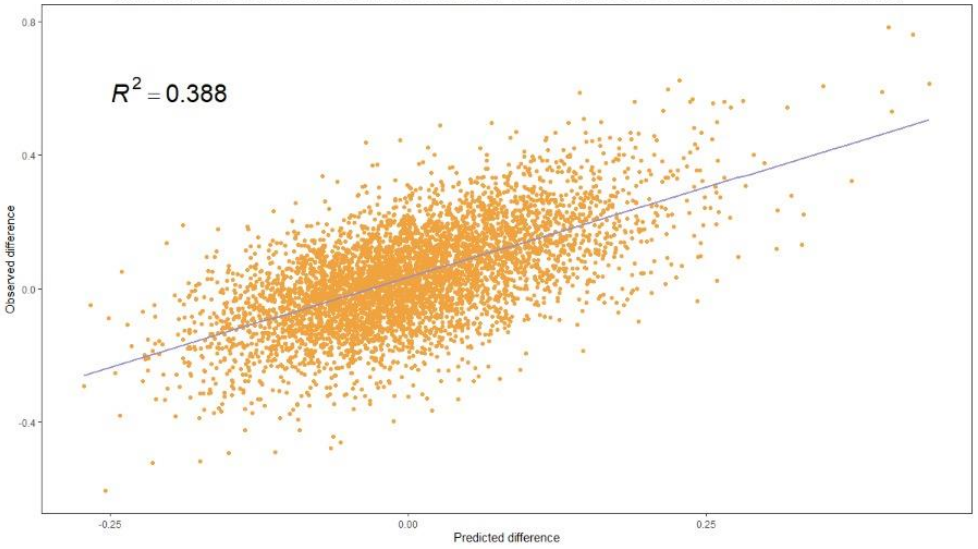
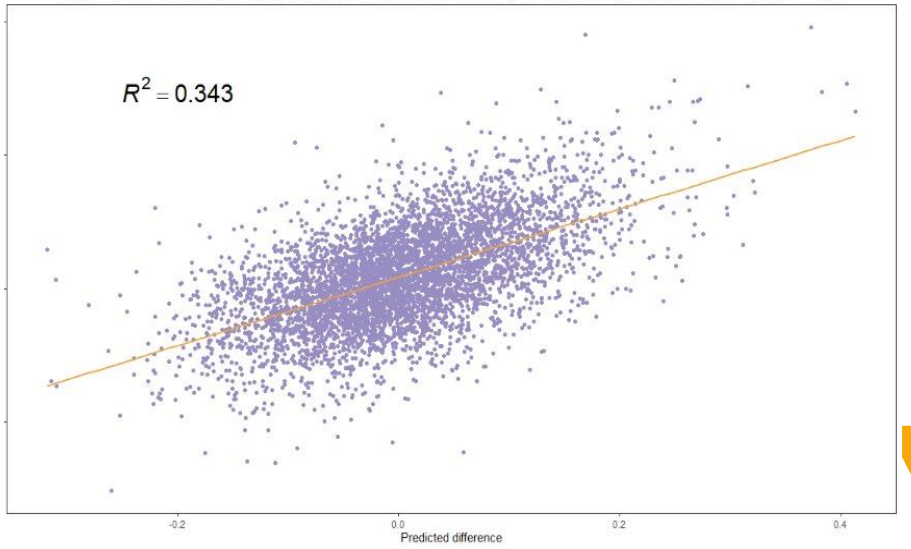


Figure 9: Observed and Predicted Difference in prop 2F Score for year 2016-2017 and 2017-2018 (Predicitive Model)



The Future, Observed and Predicted 2F Scores (year 2018 -2019)

- Relation between predicted proportion 2F scores 2016-2017 and observed scores 2018-2019
- Look at predictive value of models over 2 years
- Same 3 models

Figure 11: Observed and Predicted prop 2F score for year 2018-2019 (Model with 4 Variables)

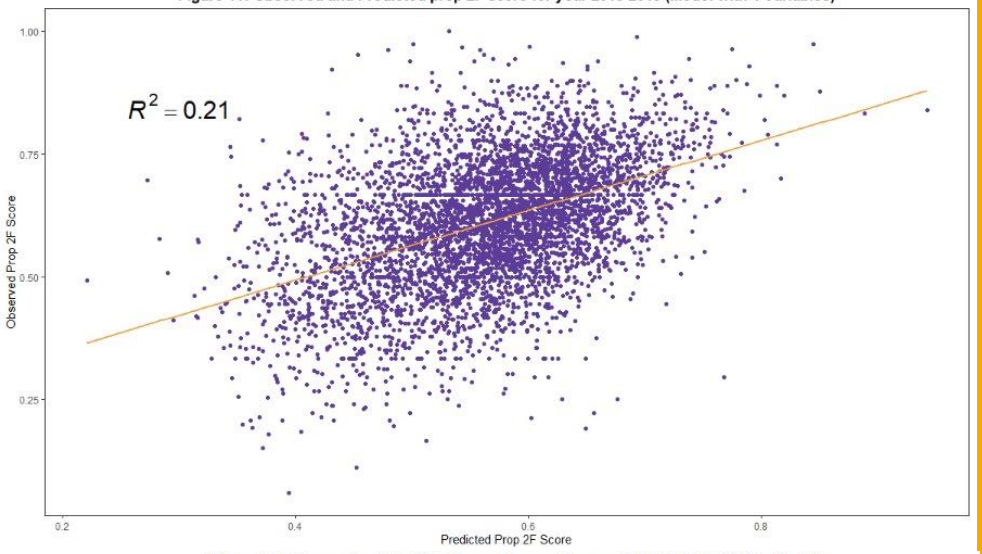


Figure 10: Observed and Predicted prop 2F score for year 2018-2019 (Simple Model)

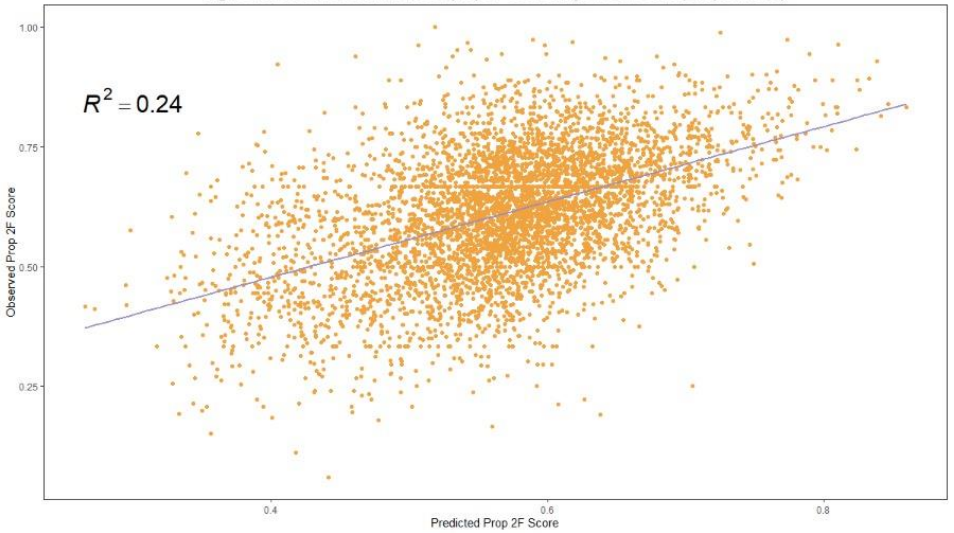
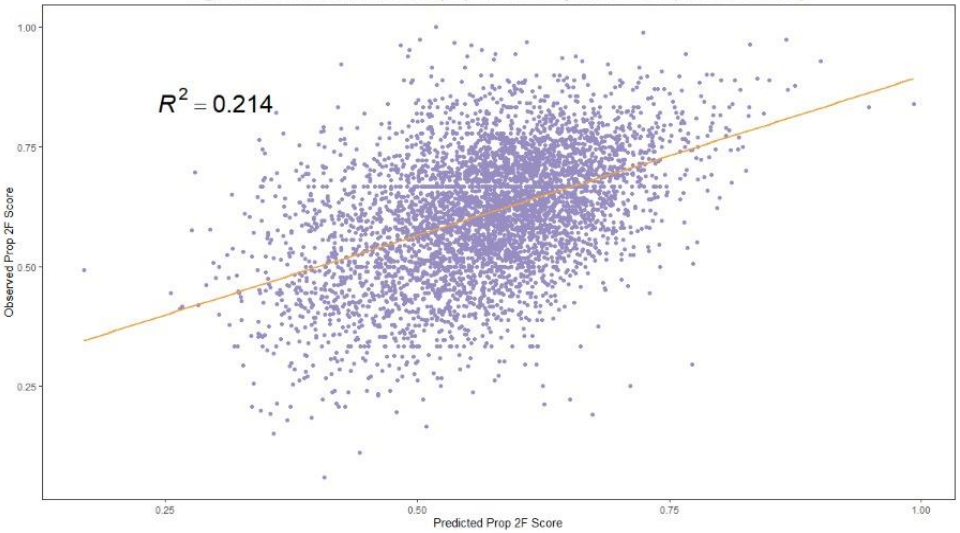


Figure 12: Observed and Predicted prop 2F score for year 2018-2019 (Predicitive Model)



Our Perception

What are good indicators?

- Omitted all rows with (left about 88%)
- Effect of variables in cohort 2015/2016 on the prop_2F score of 2016/2017.
- An indicator should have a stand alone good relationship with the outcome variable
- A low correlation is a correlation between 0.3 and 0.5
  - but for these data predicting growth so of .1 or higher
  - This made is end up with 9 possible indicators.

Figure 1: Decent Correlation

Figure 3: Distribution of Prop 2F Score Among Provinces

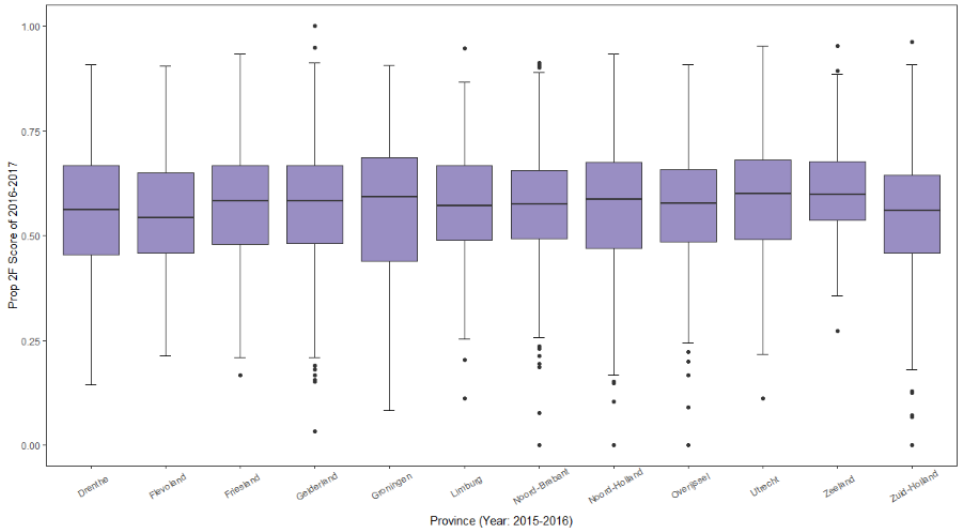
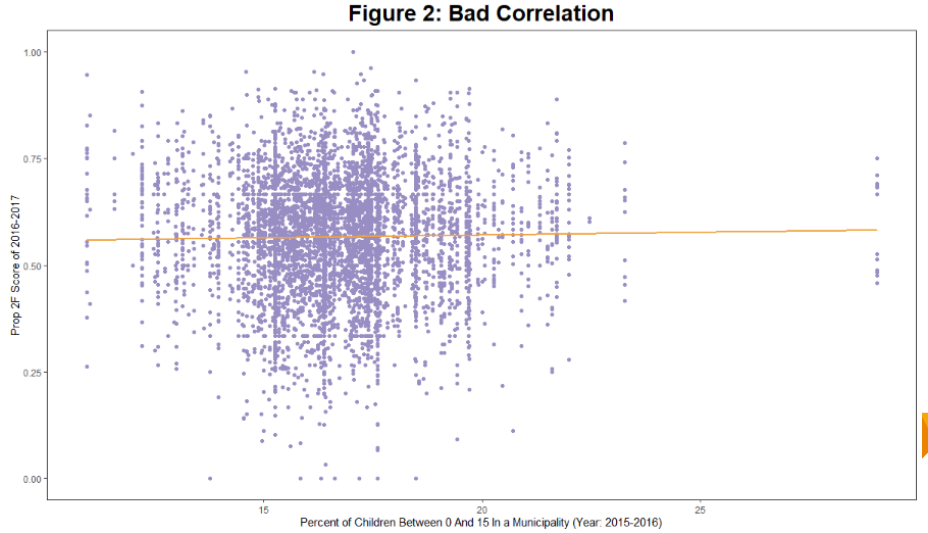


Figure 2: Bad Correlation



How Do These Indicators Perform?

Significant indicators and 3 different models

- Linear model, controlled for n_observaties.
- 9 possible indicator variables > 4 significant predictors:
- schoolweging, prop_2F (of the year before), gemeente_bevolkingsdichtheid and gemeentenummer.
- Gemeentenummer adds a lot of dummy variables > simple model
- All continuous predictors and gemeentenummer > a full predictive model
- 3 models to compare:
  - A simple indicator model with only 3 variables
  - A more complex indicator model with “4” variables
  - A fully predictive model with “12” variables.

Figure 5: Model With 4 Variables

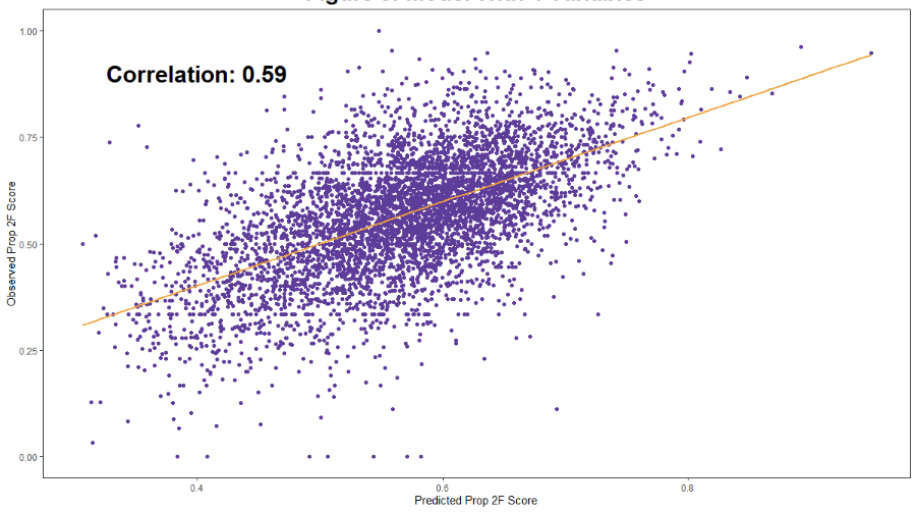


Figure 4: Simple Model

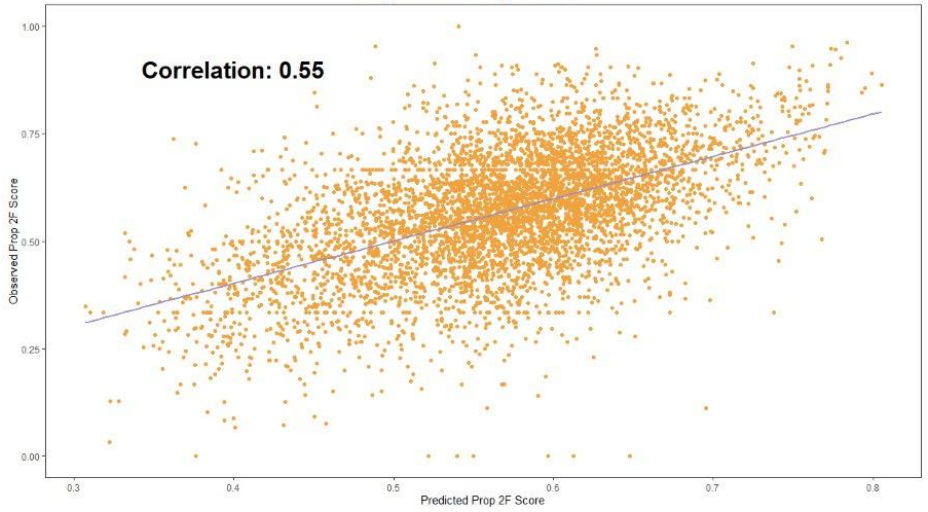
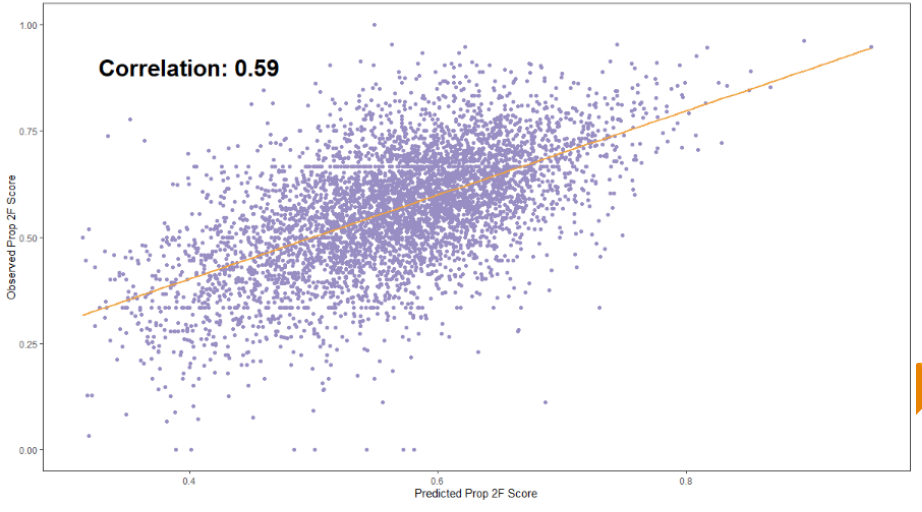


Figure 6: Model With all Predictive Variables



Linear models versus polynomial models

- Polynomial models versus linear models
- In table 1 > adjusted r squared is not that different
- In Formula 1 > variance explained by the predictors

Table 1: Comparing the linear and polynomial model for 3 different models

	Indicator low	Indicator high	Predictive
Adj.r_ln	.346	.360	.361
Adj.r_poly	.348	.359	.362

Formula 1: Linear model Indicator low

```
Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                       0.566504   0.001528 370.662  < 2e-16 ***
scale(schoolweging.x)            -0.052372   0.001661 -31.531  < 2e-16 ***
scale(prop_2F.x)                  0.039219   0.001798  21.812  < 2e-16 ***
scale(gemeente_bevolkingsdichtheid.x) -0.006418   0.001486  -4.318  1.6e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.018 on 5037 degrees of freedom
Multiple R-squared:  0.3465,    Adjusted R-squared:  0.3461
```

Our Conclusions and Recommendations

Insights and Prediction Tool

General Insights:

- Really difficult to predict growth
  - Most variables in the data set have a low correlation with the the proportion 2F score for the next year.
- 2 variables outperform the others by far:
  - proportion 2F score from the current year
  - schoolweging
- Still we were able to identify some additional variables who increase prediction of the dependent variable > gemeente bevolkinsdichtheid.

Model Insight:

- Simplest models performs better on the test data.
  - Both when we look 1 and 2 years ahead.

Conclusion:

- From our perspective it seems that some variables have a much higher predictive value than others and only incorporating these key variables might be sufficient to perform about as well as a predictive algorithm containing all variables

Table 2: Application of the simple model

10 worst school that will decrease the most

	School	Average Prop_2F Score	Expected Decrease
1	18WL	0.42	-0.29
2	16BZ	0.44	-0.28
3	15PX	0.41	-0.28
4	16KG	0.41	-0.26
5	17BY	0.40	-0.21
6	16MC	0.37	-0.20
7	14LQ	0.36	-0.19
8	19AF	0.45	-0.19
9	19LM	0.42	-0.17
10	10VR	0.44	-0.17

Simple model to find schools that would decrease dramatically in the next year.
- the “regression towards the men”
> we focus only on the 10% worst schools among the past 3 years (based on their prop 2F scores) and see whether they will continue performing worse in the upcoming year.