Current Approaches to Modeling Educational Effectiveness in Ontario:
Measuring School Performance in TDSB

TDSB Organizational Development Department

Prepared by Erhan Sinay

January, 2008

From School Effectiveness and School Improvement to Effective School Improvement:

- School Effectiveness and School Improvement have different origins: School Effectiveness measures what factors are deemed to work and why ; School Improvement is practice and policy oriented and is intended to change education in the desired direction.
- The School Effectiveness and the School Improvement paradigm can be linked in view of promoting quality improvement of education.
- In merging both approaches, here we will discuss the current use of "Modelling School Effectiveness in TDSB".

Relationship Between Performance and Context

- Without any knowledge of the context of the school, raw examination results can be misleading (Nuttal *et al*., 1989; Sammons *et al.,* 1993; Mortimore *et al.* 1994; Thomas and Goldstein, 1995; Thomas and Mortimore, 1996; Goldstein and Spiegelhalter, 1996; Gray *et al*., 1996).

- Therefore, information about school's raw examination results, such as reports published by EQAO, will always be an inadequate measure of performance.

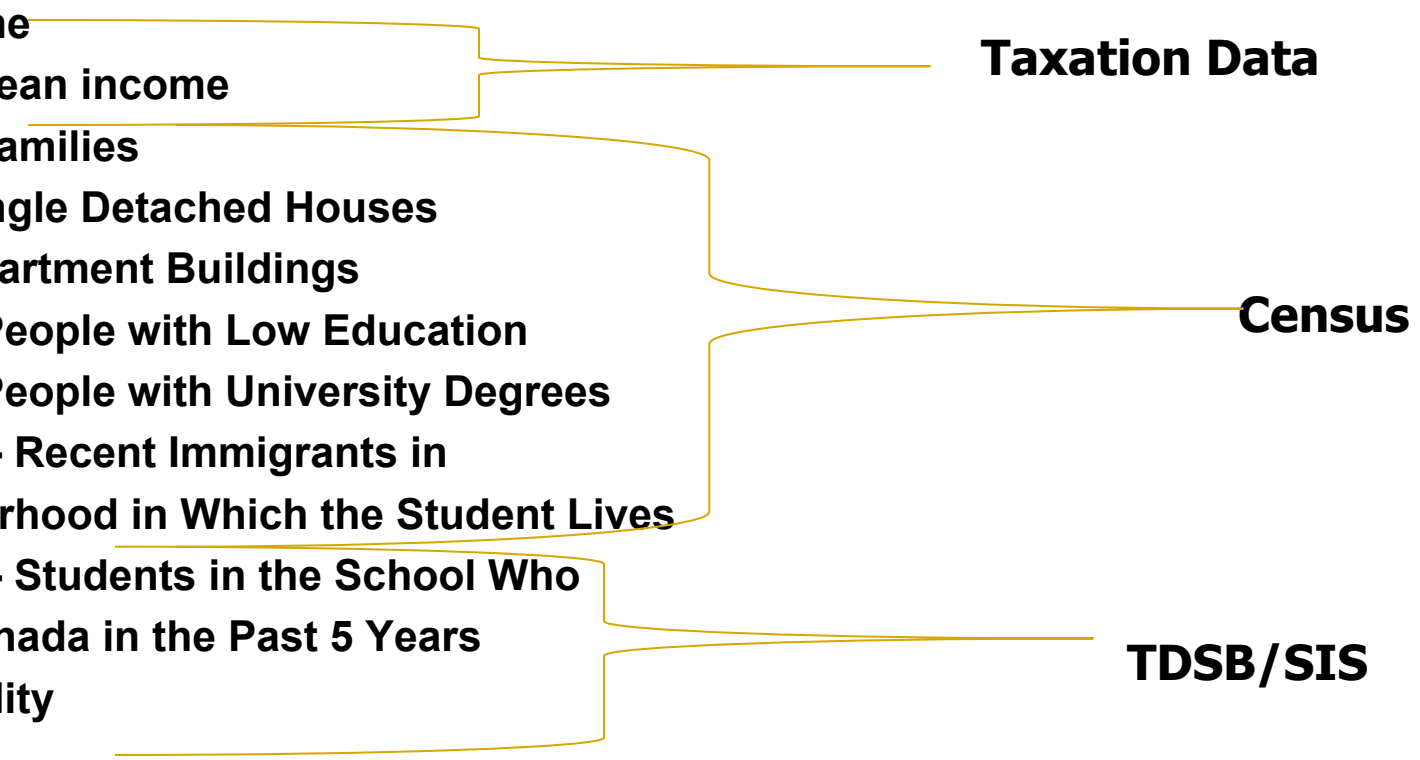
Measuring School Performance: The Learning Opportunities Index (LOI):

- Since the 1960s, the City of Toronto school boards have allocated funding to schools based on social and economic characteristics of their school Populations.

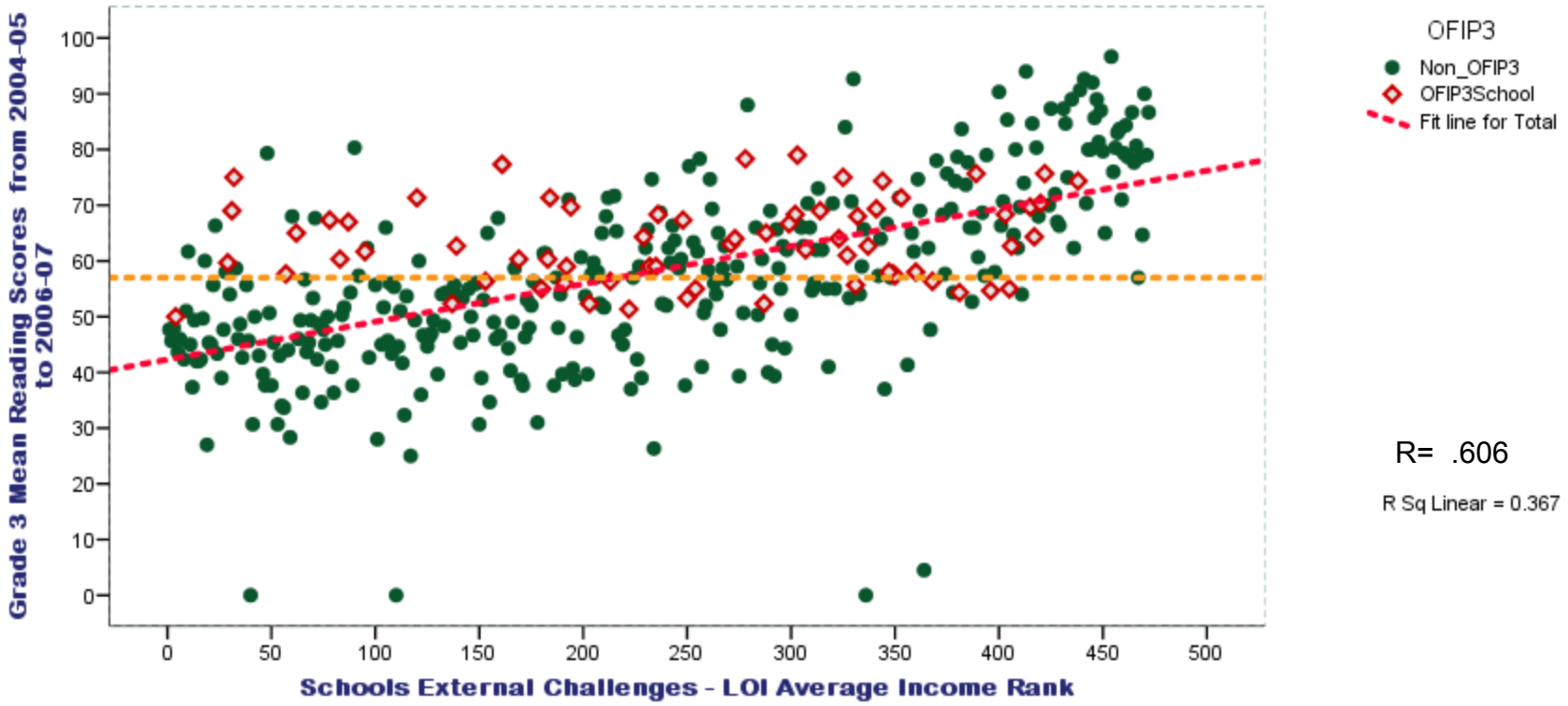
- The LOI ranks each school from most needy to least needy. These rankings are based on a range of variables or indicators that represent the multiple challenges facing the schools.

Contextual Factors Used in LOI

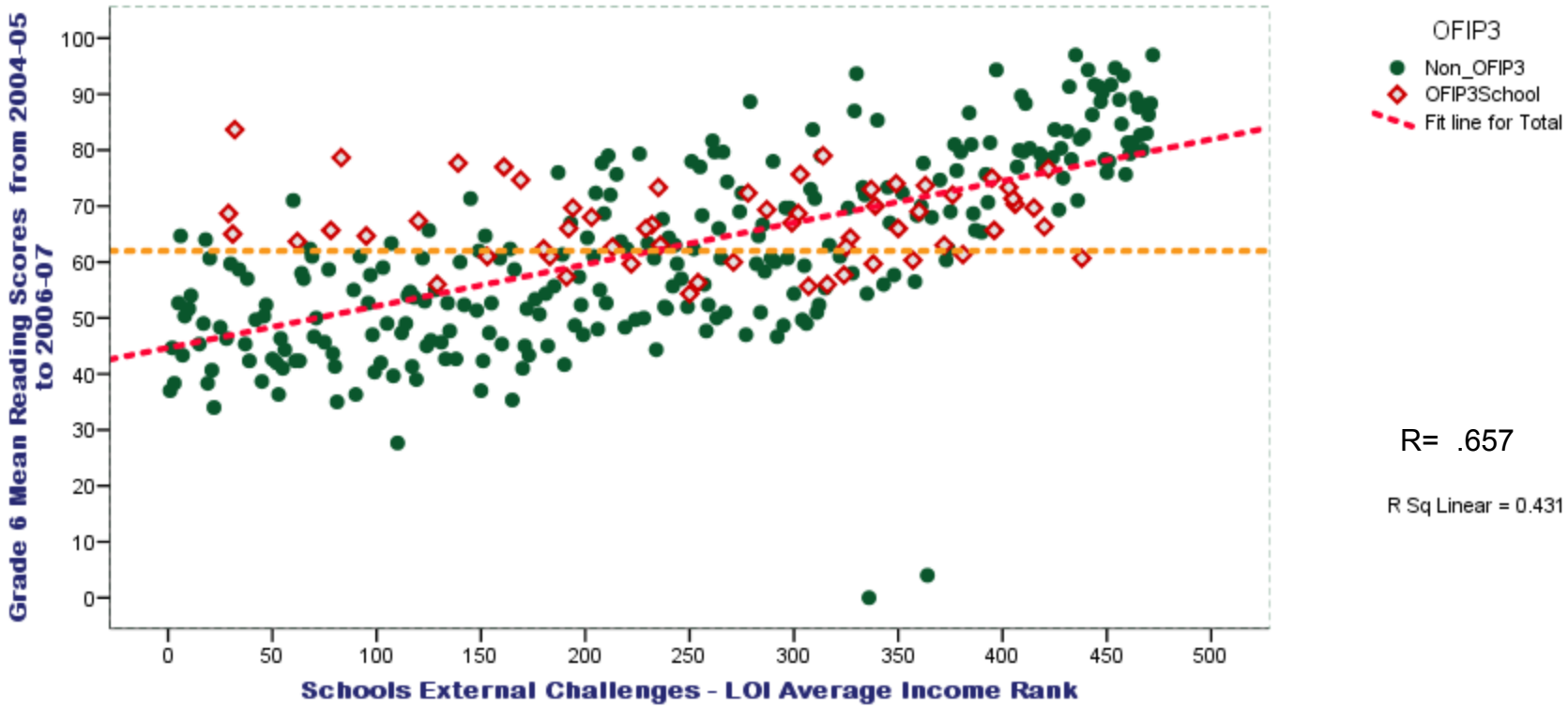
- Median Income
- Average or mean income
- Lone-parent families
- Housing – Single Detached Houses
- Housing – Apartment Buildings
- Education – People with Low Education
- Education – People with University Degrees
- Immigration – Recent Immigrants in the Neighbourhood in Which the Student Lives
- Immigration – Students in the School Who Arrived in Canada in the Past 5 Years
- Student Mobility



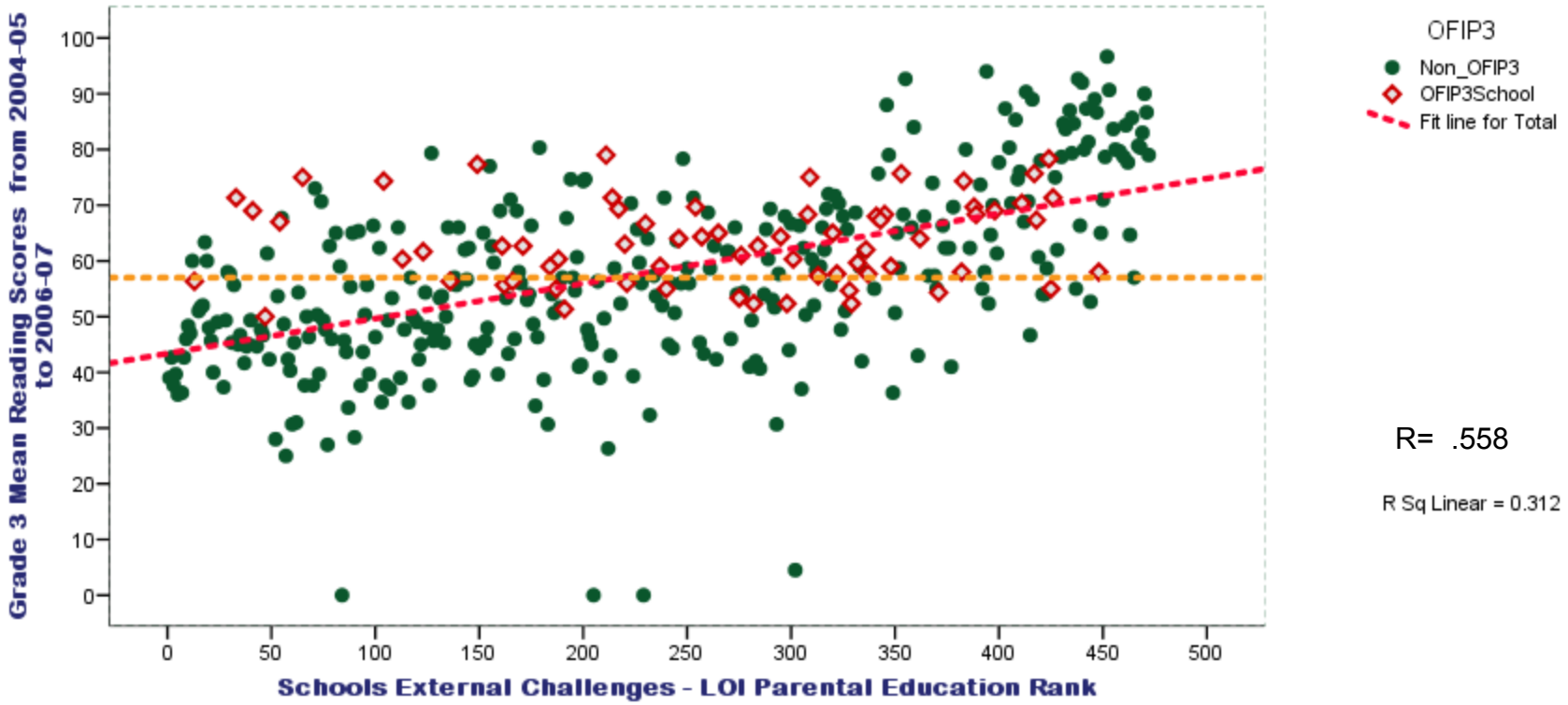
Relationship Between Performance and Context – LOI Average Income -Primary



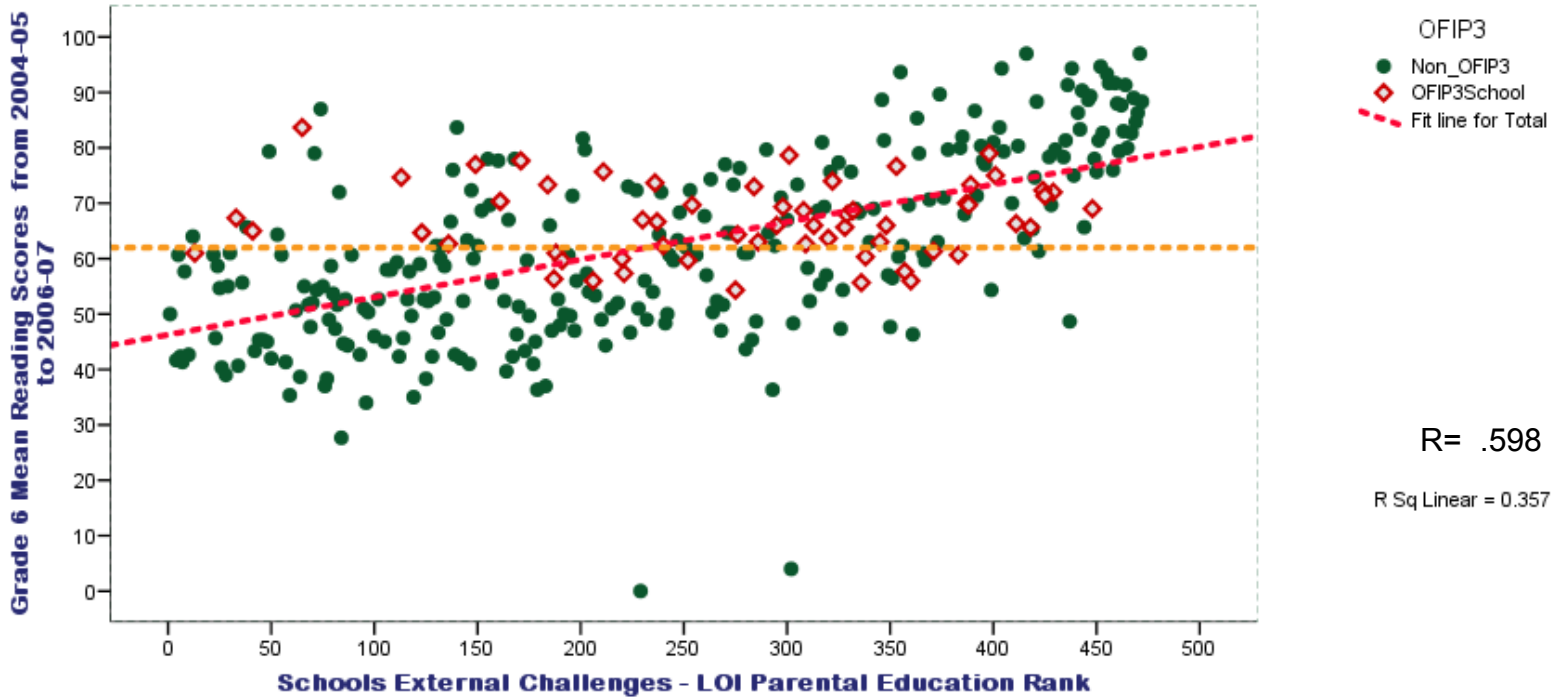
Relationship Between Performance and Context – LOI Average Income – Junior



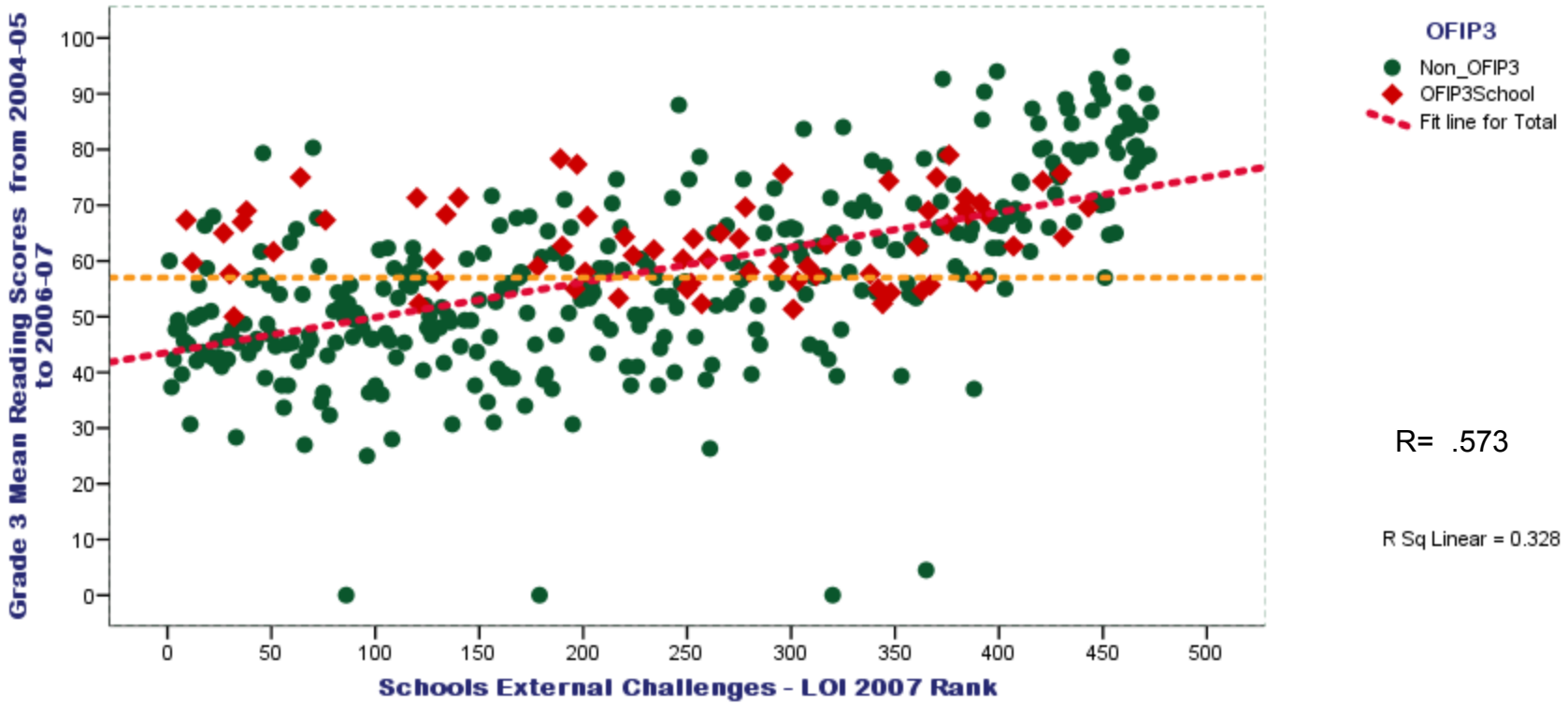
Relationship Between Performance and Context – LOI Parental Education- Primary



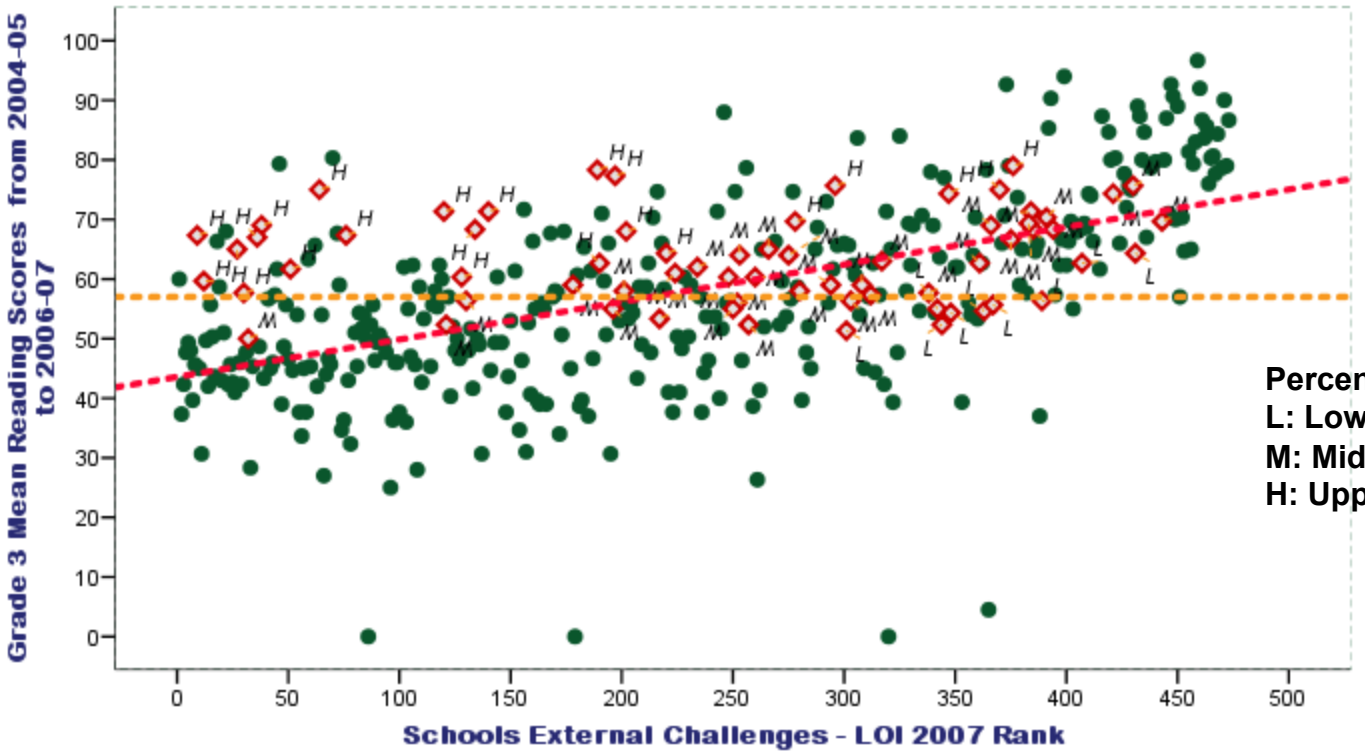
Relationship Between Performance and Context – LOI Parental Education – Junior



Learning Opportunities Index (LOI) vs. Average EQAO Performance - Primary

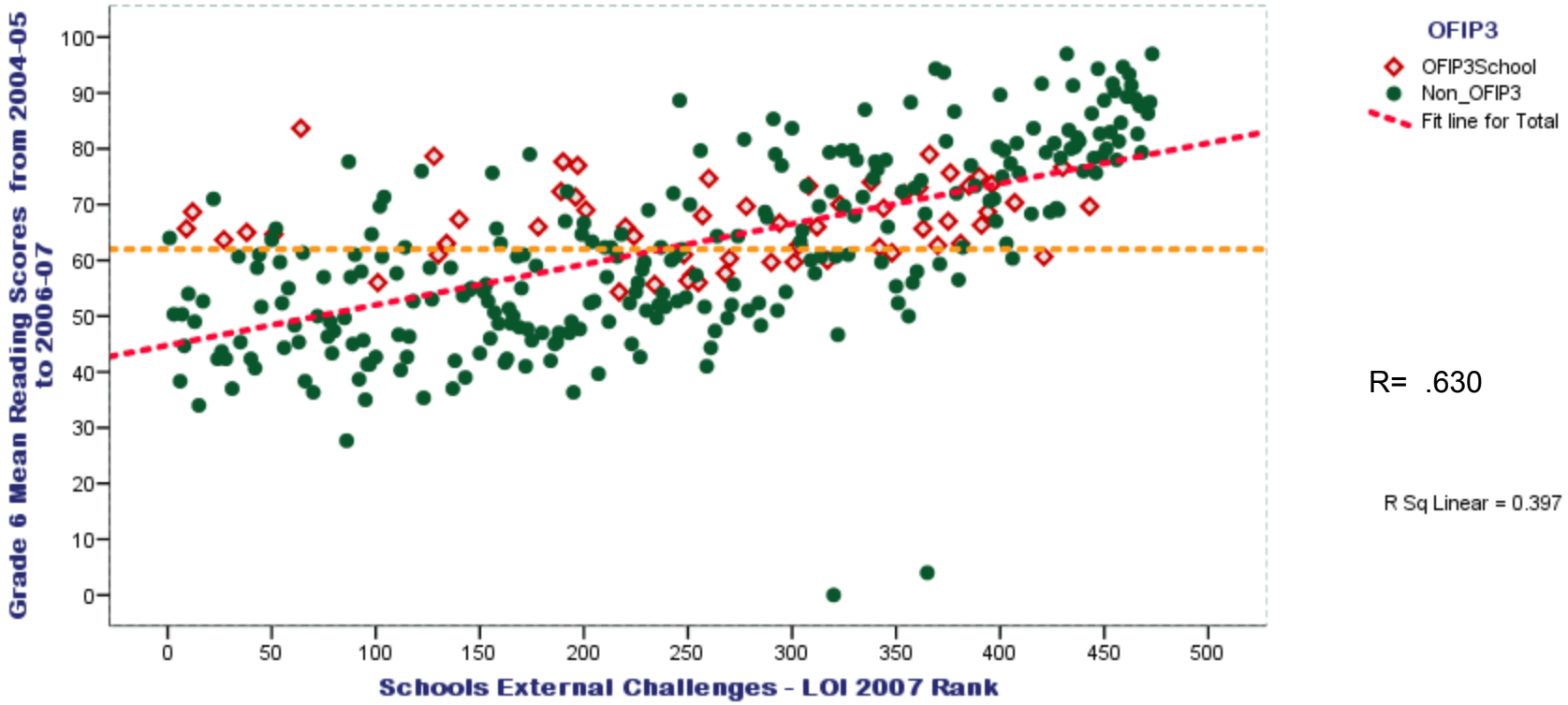


Learning Opportunities Index (LOI) vs. Average EQAO Performance - Primary

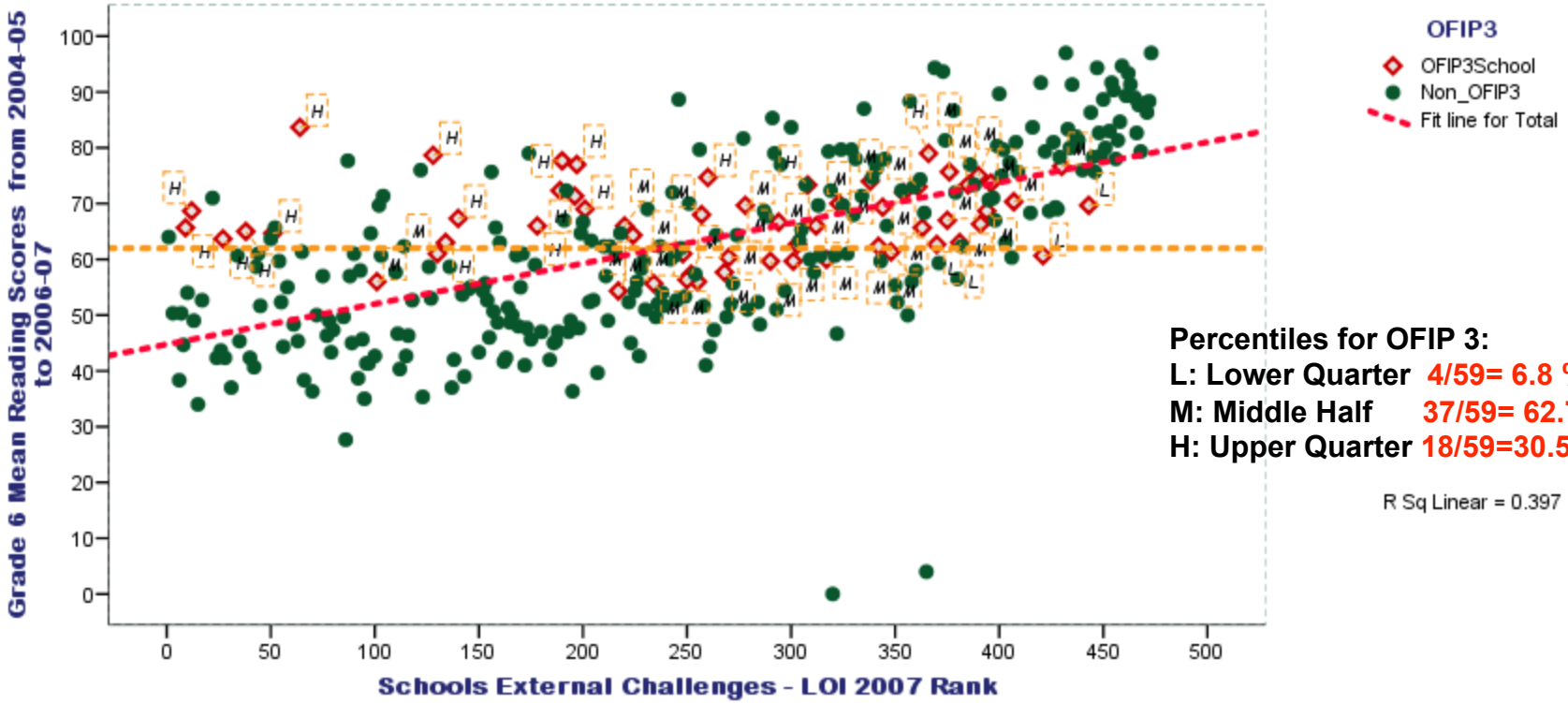


Percentiles for OFIP 3:
L: Lower Quarter 9/67=13.4 %
M: Middle Half 35/67= 52.2 %
H: Upper Quarter 23/67=34.3%

Learning Opportunities Index (LOI) vs. Average EQAO Performance - Junior



Learning Opportunities Index (LOI) vs. Average EQAO Performance - Junior



Percentiles for OFIP 3:
L: Lower Quarter 4/59= 6.8 %
M: Middle Half 37/59= 62.7 %
H: Upper Quarter 18/59=30.5%

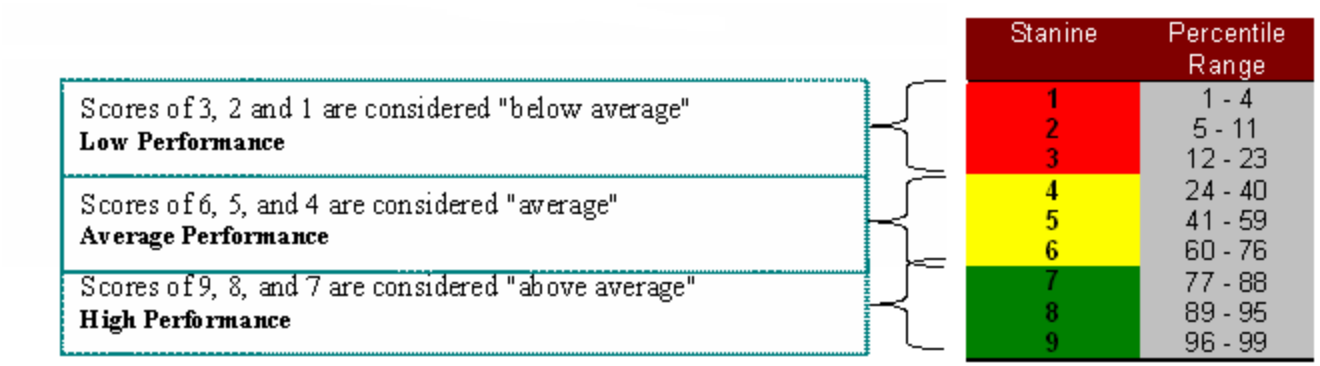
Creating a Measure of School Relative Performance-TDSB Example – Model with School Level Data

- This model is based on the school performance employing a school level longitudinal database. The scores on Grade 3 and/or Grade 6 EQAO assessments from a single school over the three years (2004-05, 2005-06, 2006-07) are transformed into a single number.

- This single score is then used as the dependent outcome, and the LOI is used as an independent variable in a linear regression model. The difference between the actual score and the predicted score is used by the model to rank schools as percentiles.

Schools Relative and Absolute Performance with Stanines

- Percentile Rank Scores are broken into nine intervals (Stanine)*, with 1 as the lowest and 9 as the highest:



Stanine	Percentile Range
1	1 - 4
2	5 - 11
3	12 - 23
4	24 - 40
5	41 - 59
6	60 - 76
7	77 - 88
8	89 - 95
9	96 - 99

*A *stanine* is an abbreviation for "standard nine.",

Schools Relative and Absolute Performance Reports:

School Name	LOI 2007	FOS	Overall School Performance					Overall Performance By Grade and Subject																			
								Grade 3													Grade							
								% L3 &L4 2005-07			Relative Percentile Rank						Overall			% L3 &L4 2005-07			Relative Percer				
											Reading		Writing		Math									Reading		Writin	
			% L3 &L4 2005-07	z-score	Absolute Percentile Rank	Overall Relative Percentile Rank	Overall Stanine	Reading	Writing	Math	%Rank	Stanine	%Rank	Stanine	%Rank	Stanine	Absolute Percentile Rank	Overall Relative Percentile Rank	Overall Stanine	Reading	Writing	Math	%Rank	Stanine	%Rank	
B PS	461	X	88.4	1.87	97	81	7	86.7	87.7	90.7	81	7	85	7	71	6	96	80	7	89.3	84.0	92.0	75	6	74	6
A PS	473	X	92.6	2.16	98	89	8	86.7	87.3	98.7	81	7	84	7	87	7	97	85	7	97.0	92.0	94.0	91	8	90	8
D PS	459	X	94.3	2.28	99	92	8	96.7	96.0	94.0	96	9	95	8	79	7	99	92	8	94.7	92.3	92.0	87	7	91	8
E PS	401	X	77.1	1.08	86	63	6	66.3	74.0	89.3	37	4	60	6	80	7	83	61	6	75.0	77.0	80.7	48	5	68	6
H PS	450	X	90.0	1.98	98	86	7	89.0	81.0	92.0	87	7	71	6	75	6	95	79	7	88.7	91.3	98.0	75	6	90	8
R PS	435	X	86.7	1.75	96	81	7	84.7	79.7	92.0	81	7	70	6	78	7	94	77	7	91.3	81.7	90.7	85	7	72	6
K PS	462	X	85.9	1.69	95	75	6	83.7	81.0	84.3	74	6	70	6	54	5	92	66	6	93.3	81.0	92.0	85	7	65	6

Low Performance

Average Performance

High Performance

Summary Next Steps:

- Most current ratings of school improvement compare the school's test score in one year with that school's score in the previous year. Differences may simply reflect variations in two student populations rather than a change in school performance.
- The fact that a school may have declined 1-2% does not take into account cohort variation. Even a group of 16-30 students can easily show cohort variation.
- For example, if a school had a Grade 3 cohort with 16 students, and 12 of the 16 achieved at Level 3/4 in 2005-6 this would be a rate of 75%. If the cohort in the school in the next year had 17 students, and again, 12 students achieved Level 3/4, this would show only 70.6%. It would appear to be a decline of almost 5%. But the reality is that there would be no real change in the school.

Summary Next Steps:
(con'd)

- The analysis of individual student data allows a sensitive examination of all aspects of a school's value added performance. Some schools that may appear to be so effective for all students in terms of the overall value added measure may not be so effective, when individual departments or different groups of students or different periods of time are examined in more detail.

- In our current school effectiveness studies, we are using individual student level data as well as school level data. A sample of HLM analysis we've done was presented at the 2007 European Association for Research on Learning and Instruction (EARLI) Conference.

Exploring the School Effects Using HLM

Model Building:

School effects were examined following a three-step procedure: unconditional model, random coefficient model and intercept-and-slopes-as-outcomes model.

Model I: Unconditional Model

- The unconditional model did not include any predictors at either Level 1 or Level 2. It is random-effect model, as school effects are defined as random across schools.
- Expressed symbolically, the unconditional model is:

LEVEL 1 MODEL

$$\text{VAR } OS_{ii} = \beta_{0j} + r_{ij}$$

LEVEL 2 MODEL

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

Table 1- Results from the Unconditional Model

The outcome variable is VAR_OS

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	0.801401	0.021931	36.542	85	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	0.19584	0.03835	85	1363.81435	0.000
level-1,	R	0.32714	0.10702			

Statistics for current covariance components model

Deviance = 4723.731507
Number of estimated parameters = 2

Results from the Unconditional Model

- The maximum likelihood estimate of within-school variance (δ^2) was 0.107 and that between-school variance (τ) was 0.038. The intra-school correlation (ρ), represented by the ratio of between-school variance to total variance, was:

  $\rho = \tau/(\tau + \delta^2) = 0.038/(0.038+0.107) = 0.26$

- Therefore, 26% of the variance in student achievement lies in between-school differences, and consequently, 74% of it lies in within-school variability.

Model III: Random Coefficient Model

- The random coefficient model regresses the OSSLT outcome measure on student characteristics. It provides estimates of variability in the regression coefficients, including the intercept and slopes, across schools. In this model, regression coefficients are assumed to vary across the population of schools.

LEVEL 1 MODEL

$$VAR\_OS_{ij} = \beta_{0j} + \beta_{1j}(VAR\_GR3_{ij}) + \beta_{2j}(VAR\_INC_{ij}) + \beta_{3j}(RVAR\_AR_{ij}) + \beta_{4j}(RVAR\_ACA_{ij}) + r_{ij}$$

LEVEL 2 MODEL

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40} + u_{4j}$$

$$VAR\_OS_{ij} = \gamma_{0c} + \gamma_{1c}*VAR\_GR3_{ij} + \gamma_{2c}*VAR\_INC_{ij} + \gamma_{3c}*RVAR\_AR_{ij} + \gamma_{4c}*RVAR\_ACA_{ij} + u_{0j} + u_{1j}*VAR\_GR3_{ij} + u_{2j}*VAR\_INC_{ij} + u_{3j}*RVAR\_AR_{ij} + u_{4j}*RVAR\_ACA_{ij} + r_{ij}$$

Table 2-Results from the Random Coefficient Model

The outcome variable is VAR_OS

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	0.898008	0.009251	97.067	84	0.000
For VAR_GR3 slope, B1					
INTRCPT2, G10	0.100487	0.009836	10.216	84	0.000
For VAR_INC slope, B2					
INTRCPT2, G20	-0.022457	0.007245	-3.099	84	0.003
For RVAR_AR slope, B3					
INTRCPT2, G30	-0.209134	0.017684	-11.826	84	0.000
For RVAR_ACA slope, B4					
INTRCPT2, G40	-0.238113	0.018896	-12.601	84	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	0.05320	0.00283	70	96.48309	0.020
VAR_GR3 slope,	U1	0.05641	0.00318	70	101.05010	0.009
VAR_INC slope,	U2	0.02607	0.00068	70	75.32743	0.310
RVAR_AR slope,	U3	0.12988	0.01687	70	222.27868	0.000
RVAR_ACA slope,	U4	0.13774	0.01897	70	174.05197	0.000
level-1,	R	0.28040	0.07862			

Note: The chi-square statistics reported above are based on only 71 of 85 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 2463.476148
Number of estimated parameters = 16

Results from Random Coefficient Model:

The results of Model II are presented in Table 2:

- The prior composite achievement in Grade 3 (Var_GR3 γ_{10} =.100 p= .000) correlates positively with OSSLT achievement within a school. That is, student OSSLT achievement increased, on average, with .100 for one point increase on the OSSLT outcome.
- The average Income gap (Var_INC) was -.022, indicating that students in low income neighborhood scored on average .022 points lower than other students, net of previous achievement, at risk in Grade 10 and program of study effects.
- The at risk in Grade 10 RVAR_AR (γ_{30} = -.209 p= .000) indicates that students at risk in Grade 10 scored on average .209 points lower than other students.
- Students who were not in Academic program RVAR_ACA (γ_{40} =-.238 p= .000) tends to score significantly lower than the other students.

Results from Random Coefficient Model (con'd) :

- Random effects provide estimates of variance and test the hypothesis that the variance was zero across schools:
- The relationship between prior achievement (VAR_GR3) and OSSLT; At-risk in Grade 10 (RVAR_AR) and OSSLT and Student Program of Study with OSSLT varied significantly across the schools
- The estimated variance for income did not vary significantly across the schools.

Model III: Intercept-and slopes-as-outcomes Model

- Model at this step becomes:

LEVEL 1 MODEL

$$VAR\_OS_{ij} = \beta_{0j} + \beta_{1j}(VAR\_GR3_{ij}) + \beta_{2j}(VAR\_INC_{ij}) + r_{ij}$$

LEVEL 2 MODEL

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(PROPESLE_j) + \gamma_{02}(PROPPROG_j) + \gamma_{03}(PROPINC_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(PROPESLE_j) + \gamma_{12}(PROPPROG_j) + \gamma_{13}(PROPINC_j) + u_{1j}$$

$$\beta_{2j} = \gamma_{20}$$

It is assumed that the intercept and the slope vary not only as a function of the three predictors, PROPESLE, PROPSPEC, and PROPPROG, but also a function of the unique school effect. Substituting the student level model
with school level equations yields the combined model:

$$VAR\_OS_{ij} = \gamma_{00} + \gamma_{01}*PROPESLE_j + \gamma_{02}*PROPPROG_j + \gamma_{03}*PROPINC_j + \gamma_{10}*VAR\_GR3_{ij} + {}_{11}* \gamma_{11}*PROPESLE_j*VAR\_GR3_{ij} + \gamma_{12}*PROPPROG_j*VAR\_GR3_{ij} + \gamma_{13}*PROPINC_j*VAR\_GR3_{ij} + \gamma_{20}*VAR\_INC_{ij} + u_{0j} + u_{1j}*VAR\_GR3_{ij} + r_{ij}$$

Table 3: Intercept-and slopes-as-outcomes model:

The outcome variable is VAR_OS

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	0.211218	0.057222	3.691	81	0.001
PROPESLE, G01	0.815429	0.098592	8.271	81	0.000
PROPPROG, G02	0.726469	0.052767	13.767	81	0.000
PROPINC, G03	-0.138514	0.068714	-2.016	81	0.047
For VAR_GR3 slope, B1					
INTRCPT2, G10	0.416456	0.058184	7.158	81	0.000
PROPESLE, G11	-0.394136	0.122414	-3.220	81	0.002
PROPPROG, G12	-0.354808	0.056196	-6.314	81	0.000
PROPINC, G13	0.154719	0.059267	2.611	81	0.011
For VAR_INC slope, B2					
INTRCPT2, G20	-0.035764	0.009374	-3.815	7358	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	0.06997	0.00490	76	198.95981	0.000
VAR_GR3 slope,	U1	0.05867	0.00344	76	123.38351	0.001
level-1,	R	0.31351	0.09829			

Note: The chi-square statistics reported above are based on only 80 of 85 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 3933.283647
Number of estimated parameters = 4

Results from Intercept-and slopes-as-outcomes model

- Results from the intercept- and slopes- as outcome model indicates a significant relationship between OSSLT outcomes and previous achievement in the following instances :

- the proportion of the students from low income neighborhoods,
- the proportion of students in English as a Second Language (ESL) and,
- the proportion of students not in the Academic Program in school was significantly related to OSSLT and previous achievement.

Summary Results and Next Steps

- The significant variation in average achievement among schools indicates that TDSB schools are heterogeneous in student achievement. The majority of variation in OSSLT achievement is within, not between, schools.

- The findings of the study show that neighborhood income (as a measure of SES), previous achievement, student program of study, and at-risk status in Grade 10 are significantly related to student achievement.

- One of the major findings of this study is that prior achievement (measured through provincial tests) is a significant predictor of current achievement. That is, in this study, students' third, sixth and ninth grades achievement in EQAO significantly predicts their tenth grade performance in OSSLT.

Summary Results and Next Steps

(con'd)

- This study investigates school effects using two-level hierarchical modeling. Variables are included independently at each level, without modeling interactions between variables. Next steps will be to incorporate these interactions into the model.

- In further analyses, we intend to expand and modify the study by exploring the similarities and differences of the student and school effects on to the other subjects, such us reading, writing and mathematics, science.

- We will also introduce more school level variables to account for the part of the variance that student background characteristics alone can not explain. This will help us identify or determine the characteristics of the effective and less effective schools, after taking their student's characteristics into considerations.