Interpolating PISA Data with the World Development Indicators

⊥ Hammed Oluwaseyi Musibau
⊥ Maria Yanotti
⊥ Joaquin Vespignani

*Tasmanian School of Business and Economics,
University of Tasmania.

AEC 2022

July 12, 2022

UNIVERSITY of TASMANIA

Outline

UNIVERSITY of TASMANIA

Background of the study

1. Education has been established in the literature as a proxy for human capital (Hanushek and Woessmann 2012, JEL)
2. Previous studies focus on quantity of education with less action on education quality.
3. Why quality of education is so important?
4. What is the stand of literature about the measure of education quality?
5. Why PISA is so important?
6. What is the main issue with PISA?
7. What is the main contribution of this study?

UNIVERSITY of TASMANIA

Background of the study

Why worry about education quality?

Cognitive Development (see, Hanuszek and Woessman (2008, 2012, JEL)

→

Empowers citizens /Outcome needed (Glewwe & Muralidharan, 2016)

→

Equity in society (Brueckner, Van, & Vespignani 2022, AE)

Review on measure of education quality

- The measure of education quality is challenging:
- Previous studies have measured quality of education using literacy rates (Hanushek and Woessman 2009), student teacher ratio Clements (2002).
- However, recent studies like Engel et al (2019), Hanushek (2015) and Hanushek and Woessman (2021) used PISA, and suggest that it is a reliable measure of education quality.
- Why PISA is so important?

UNIVERSITY of TASMANIA

Why PISA is so important?

Recently, the international large-scale assessments (ILSA) have been on the rise, with the Programme for International Student Assessment (PISA) seen as a reliable measure of education quality in term of outcome:

- The PISA is a dataset on education outcome launched in 1999 by OECD,
- It is a tri-annual study that measures student knowledge and skills in science, math and reading literacy in 2000, 2003, 2006,2009, 2012 and 2015.
- Assessing whether 15-year-olds are able to apply what they have learned in school in real life situations.
- Data from PISA are used in education policy formation in many countries (Hopfenbeck et al., 2018).
- What is the main issue with PISA?

Aims and main contribution

- Our main contribution is to extend the PISA data on education outcomes collected every three years into a yearly series for 37 OECD countries.

- Secondly, unlike the previous studies, we use of factor modelling and Chow-Lin 1971 to interpolate the PISA data using the World bank development Indicators. Factor modeling and Chow-Lin allow us to maintain similar trend between Chow Lin output and the original PISA data.

UNIVERSITY of TASMANIA

Summary of empirical studies employing mixed frequency approach.

S/N	Research	Sample	Time	Approach	Method	Findings
1	Rashid and Jehan (2013)	Pakistan	1971 to 2010	Mixed (annual GDP, investment spending, and government expenditure to quarter)	Chow Lin (1971)	Chow Lin output and annual data exbits similar trend
2	Chikamatsu et al (2018)	Japan	2003 to 2014	Mixed Frequency(annual to quarter)	OLS, Chow Lin, Denton, Fernandez and Litterman Models	The disaggegating models that used the timely monthly supply-side indicators serve as efficient tools for predicting GDP growth rates
3	Di Giacinto et al (2021)	Italy	2014 - 2017	Mixed Frequency Data (annual unemployment rates into monthly series)	Chow Lin (1971)	The newly quarterly index is consistent with the official annual GDP data
4	Petrosky nadeau and Zhang (2021)	USA	1948–2017	Mixed Frequency (The annual unemployment rates into monthly series)	Denton 1971	The newly high frequency series are quantitatively consistent with that of the historical umeployment series

Data and sample

- 37 OECD countries from 2000 to 2015
- Our dependent variable comprises PISA's three core competencies: mathematics, reading, and science (tri- annual %) – Hanuszek (2012), JEL).
- Our independent variables comprise all the 22 variables available on the proxies of education quality from the World Development Indicators (WDI, 2020)[2], which we further classified into the main three key components of pupil-teacher ratio, education compensation, and education expenditure (annual).

UNIVERSITY of TASMANIA

[2] https://databank.worldbank.org/source/world-development-indicators

Data and Sample

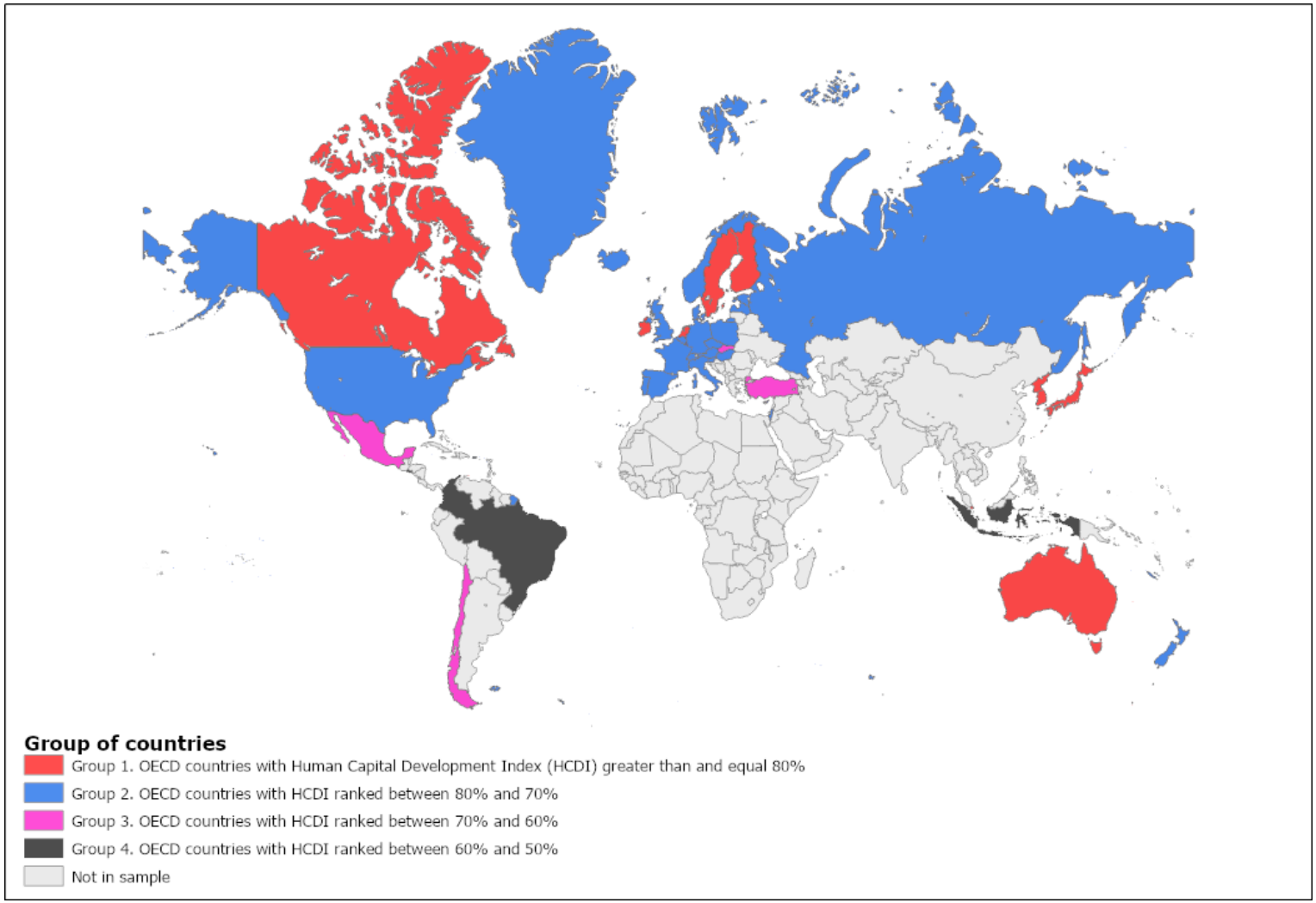
- In this section, we used the annual data on the proxies of education quality from World Development Indicators (WDI) database to interpolate the PISA data retrieved from OECD database.
- We gathered data for all the 37 OECD countries that have a complete PISA data from 2000 to 2015.
- We grouped the 37 OECD countries into 4 groups using the Human Capital development index (HCDI) to deal with the missing data in the annual observation of World development Indicators (WDI):

UNIVERSITY of TASMANIA

Data and Sample

- Group (1): OECD countries with HCDI $\geq$ 80% and above (i.e. Singapore, Japan, Korea, Rep., Hong Kong, Finland, Ireland, Australia, Sweden, Netherlands and Canada);
- Group(2): OECD countries with HCDI rank between 80% and 70% (i.e. Germany, Austria, Slovenia, Czech Republic, United Kingdom, Portugal, Denmark, Norway, Italy, Switzerland, New Zealand, France, Israel, United States, Macao SAR, China, Belgium, Estonia, Poland, Spain, Iceland, Russian Federation, Latvia, Hungary and Luxembourg);
- Group (3): OECD countries with HCDI rank between 70% and 60% (i.e. Slovak Republic, Luxembourg, Turkey , Chile and Mexico);
- Group (4): OECD countries with HCDI rank between 60% and 50% (i.e. Colombia, Peru, Brazil, Indonesia, El Salvador).

Figure: OECD Countries Classification by Human Capital Development Index



UNIVERSITY of TASMANIA

Methods

In this section, we outline the econometrics procedures we used to interpolate PISA data on education outcomes for 37 OECD countries. To fill these missing observations in PISA data from tri-annual to annual series, we followed two-steps econometrics procedures:

- First Step, we build factors to capture all the available annual series from WDI for education quality proxies.
- Second Step, we then used factors created in the first step to perform the interpolation of PISA data using Chow Lin Methodology.

UNIVERSITY of TASMANIA

Step one - Factor Building Using The World Development Indicators:

$$R_{t,i} = [R_{it}^{PrePrim}, R_{it}^{Prim}, R_{it}^{LowSec}, R_{it}^{Sec}, R_{it}^{UpSec}, R_{it}^{Ter}]] \quad (1)$$

$$C_{t,i} = [C_{it}^{Tot}, C_{it}^{Prim}, C_{it}^{Sec}, C_{it}^{Ter}]] \quad (2)$$

$$E_{t,i} = \Big[E_{it}^{CurPrim}, E_{it}^{CurSec}, E_{it}^{CurTer}, E_{it}^{CurTot}, E_{it}^{GovPrim},$$

$$E_{it}^{GovSec}, E_{it}^{GovTer}, E_{it}^{GovTot}, E_{it}^{GovGDP}, E_{it}^{Prim},$$

$$E_{it}^{Sec}, E_{it}^{Ter}]$$
(3)

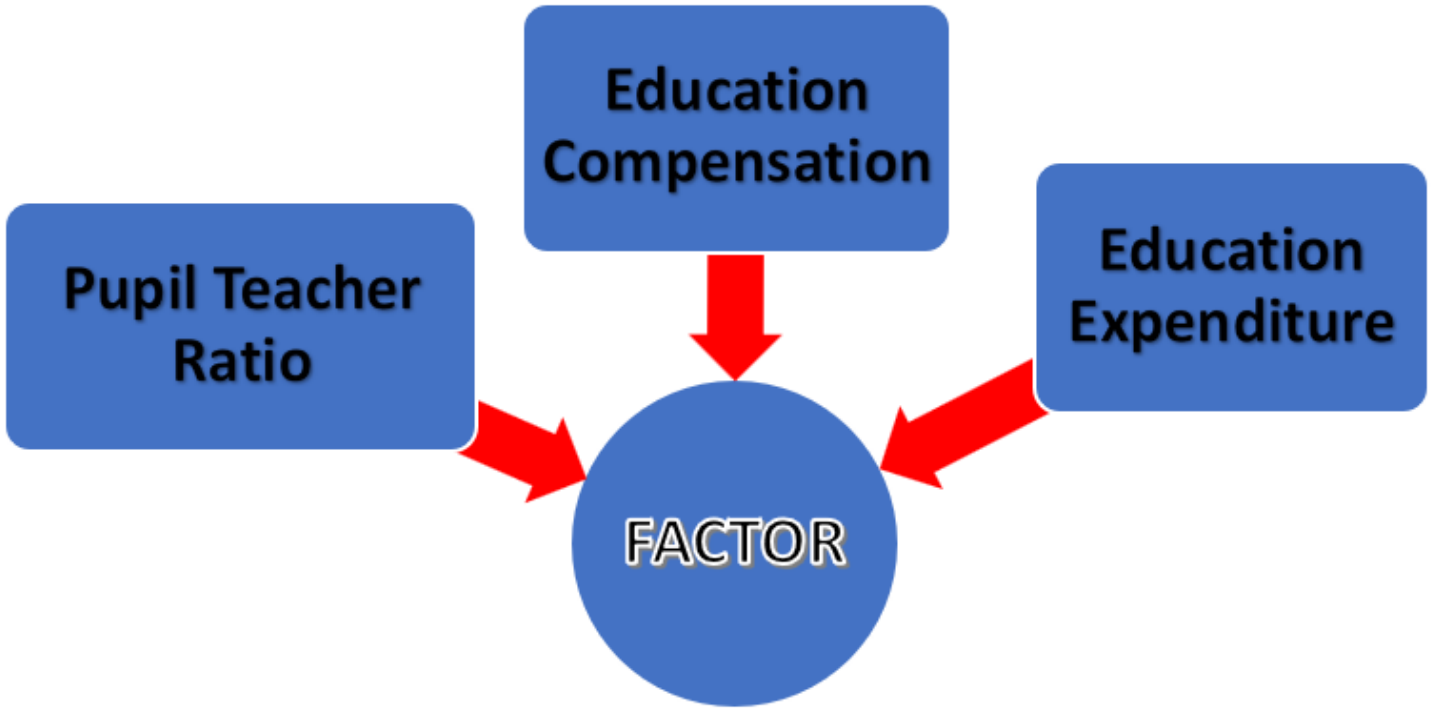
The World Bank Development Indicators

S/no	Variable	Variable
1	PupTeaRatioPrePrim	Pupil Teacher Ratio, Preprimary
2	PupTeaRatioPrim	Pupil Teacher Ratio, primary
3	PupTeaRatioLowSec	Pupil Teacher Ratio, Lower Secondary
4	PupTeaRatioSec	Pupil Teacher Ratio, Secondary
5	PupTeaRatioUpSec	Pupil Teacher Ratio, Upper Secondary
6	PupTeaRatioTer	Pupil Teacher ratio, Tertiary
7	AllEdStaffCompPrim	All Education Staff Compensation, Primary
8	AllEdStaffCompSec	All education Staff compensation, secondary
9	AllEdStaffCompTer	All education staff compensation, tertiary
10	AllEdStaffComp	All Education Staff Compensation
11	ExpPrim	Expenditure on education, primary
12	ExpSec	Expenditure perstudent, secondary
13	ExpTer	Expenditure on education, tertiary
14	GovExppsSec	Government expenditure per student, secondary
15	GovExppsPrim	Government expenditure per student, primary
16	GovExppsTer	Government expenditure per student, tertiary
17	GovExpEd	Government expenditure on ed-ucation, total
18	GovExpEdGDP	Government eXpenditure on education, as a % of GDP
19	CurrEdExpPrim	Current education expenditure, primary
20	CurrEdExpSec	Current education expenditure, secondary
21	CurrEdExpTer	Current education expenditure, tertiary
22	CurrEdExp	Current education expenditure, total

Construction of a Single Related Series

- To construct a single related series, we then calculate a weighted average of all factors for individual countries ($FACTOR = factcomp * 0.33 + factexp * 0.33 + factptr * 0.33$) to create a related series.

  Our Related Series:



UNIVERSITY of TASMANIA

Chow Lin Interpolation

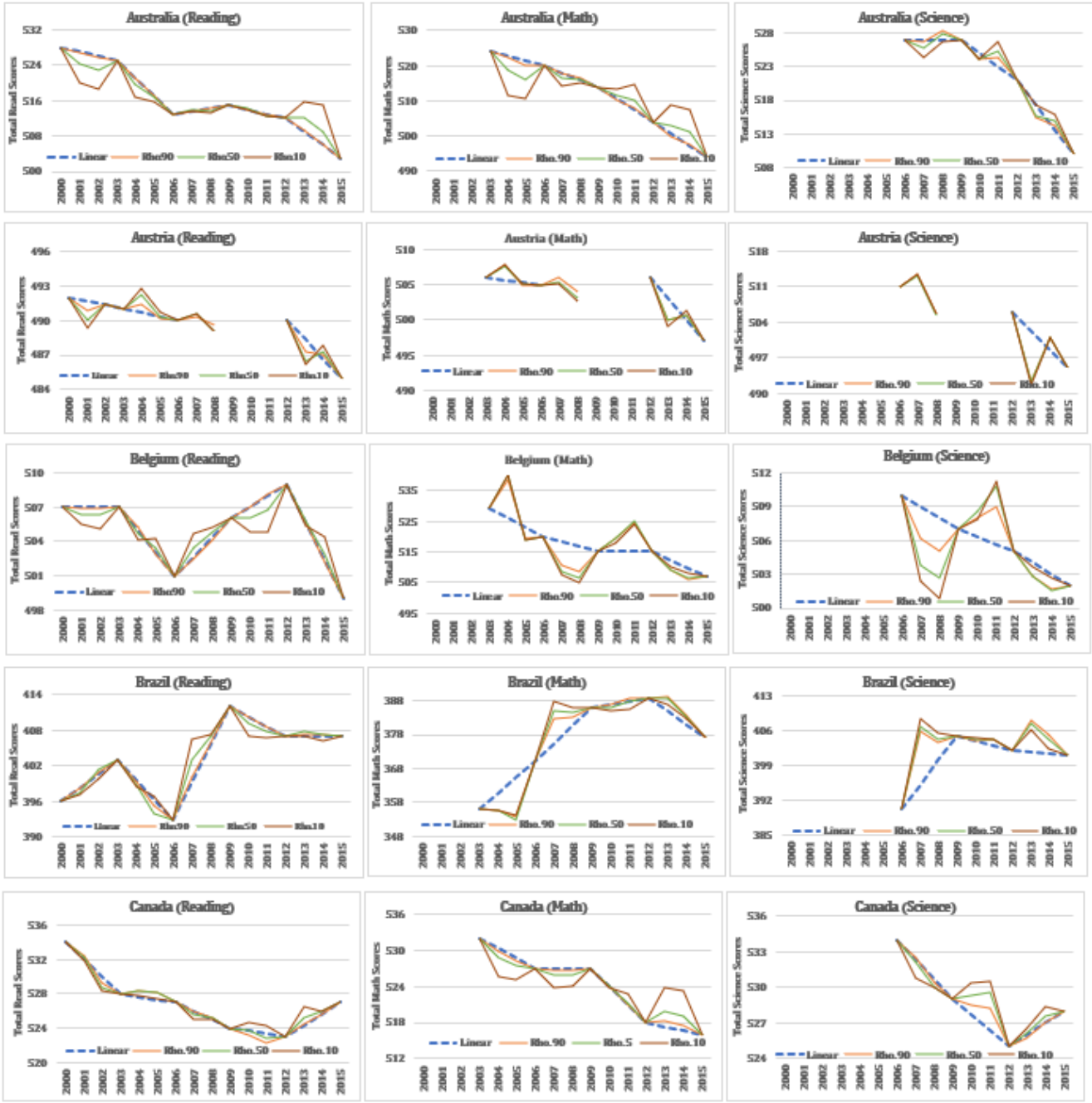
Step two - To interpolate PISA data, we used Chow-Lin (1971), a regression-based interpolation technique that relates lower-frequency indicator series ($PISA_t$), to a higher-frequency benchmark series ($FACTOR_k$), observed in every year k to obtain an annual interpolated series (X_t).

$$PISA_t^{3y} = \beta_i FACTOR_{i,k}^{y} + e_k \tag{4}$$

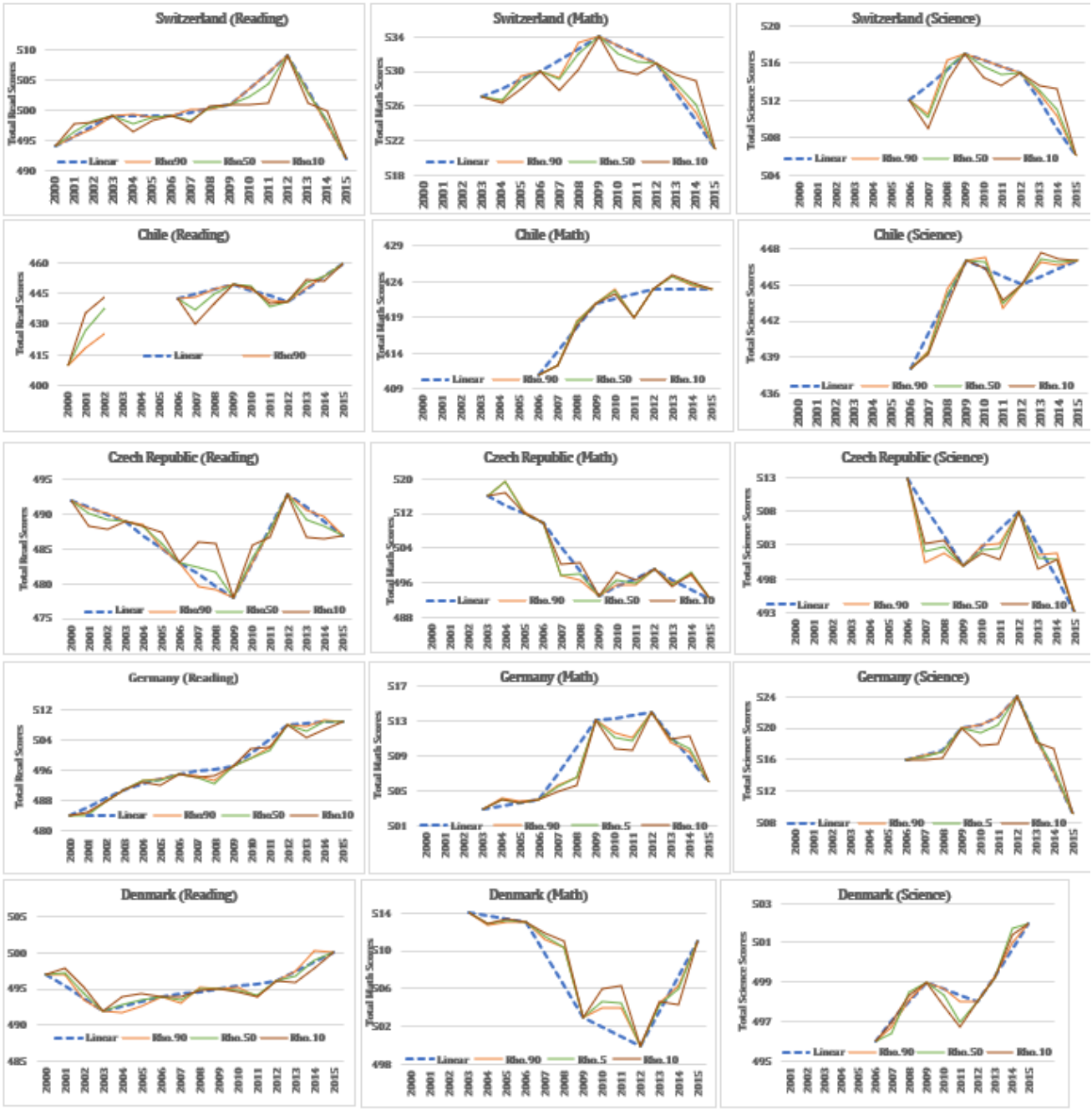
$$X_t^* = \hat{\beta}_i FACTOR_{i,k}^{y} + e_k \tag{5}$$

$$\sum_{t=1_k}^{4_k} X_t^* = FACTOR_k^{y} + e_k \tag{6}$$

where X_t contains observations for each of the years (t). This final condition ensures that the sum of the tri-annual log changes in the interpolated series (X_t), sums to the annual log change of the actual data.

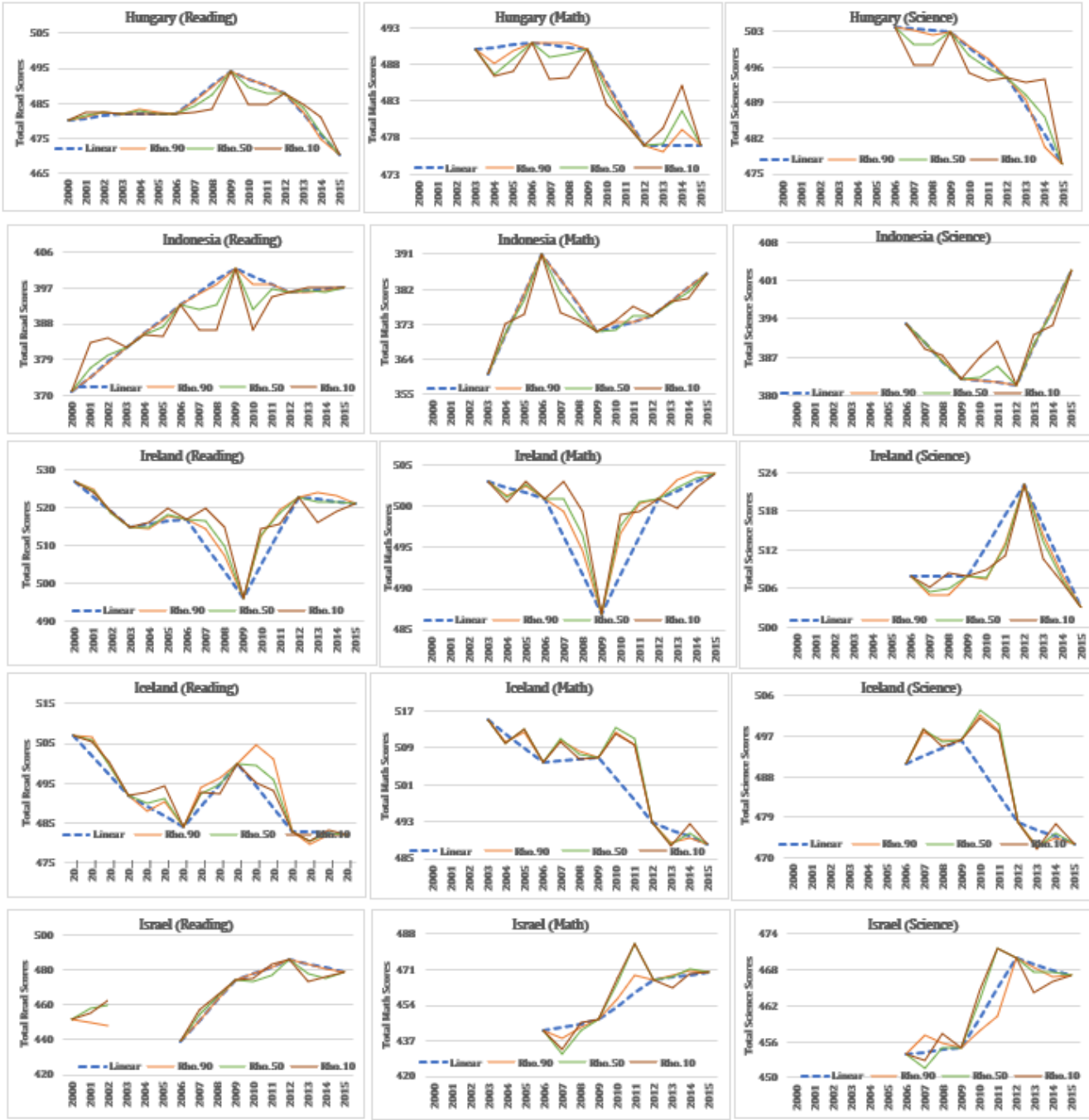


UNIVERSITY of TASMANIA

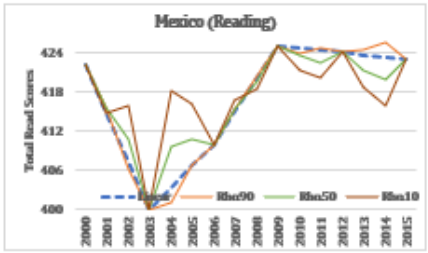


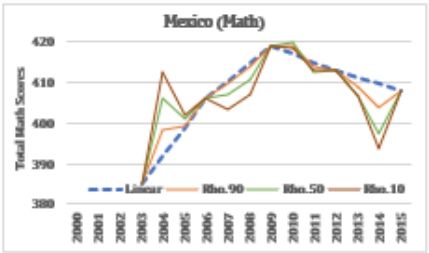
UNIVERSITY of TASMANIA

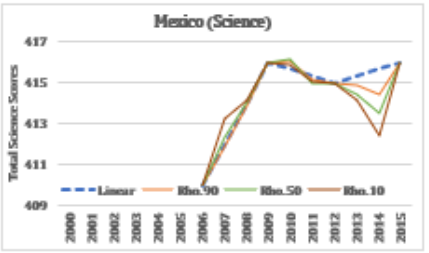

Spain (Reading)
Spain (Math)
Spain (Science)
Estonia (Reading)
Estonia (Math)
Estonia (Science)
Finland (Reading)
Finland (Math)
Finland (Science)
France (Reading)
France (Math)
France (Science)
United Kingdom (Reading)
United Kingdom (Math)
United Kingdom (Science)
Linear
Rho.90
Rho.50
Rho.10

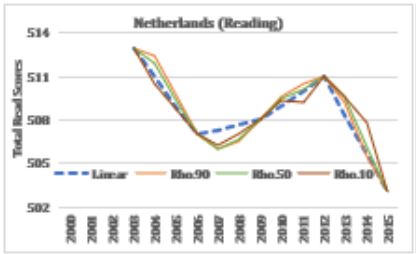

Hungary (Reading)
Hungary (Math)
Hungary (Science)
Indonesia (Reading)
Indonesia (Math)
Indonesia (Science)
Ireland (Reading)
Ireland (Math)
Ireland (Science)
Iceland (Reading)
Iceland (Math)
Iceland (Science)
Israel (Reading)
Israel (Math)
Israel (Science)
Linear
Rho.90
Rho.50
Rho.10

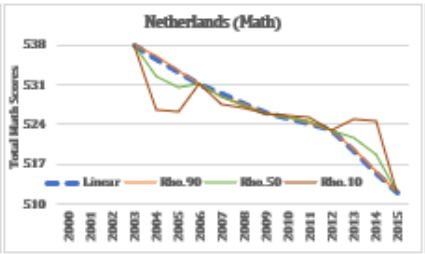

Italy (Reading)
Italy (Math)
Italy (Science)
Japan (Reading)
Japan (Math)
Japan (Science)
Korea, Rep. (Reading)
Korea, Rep. (Math)
Korea, Rep. (Science)
Luxembourg (Reading)
Luxembourg (Math)
Luxembourg (Science)
Latvia (Reading)
Latvia (Math)
Latvia (Science)
Linear
Rho.90
Rho.50
Rho.10

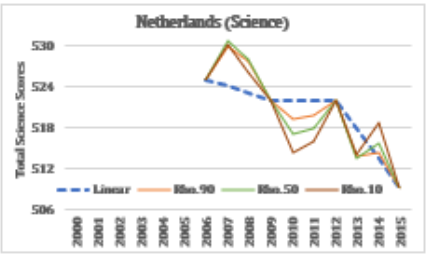


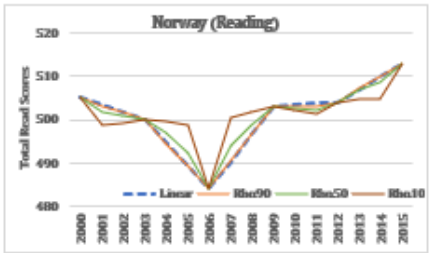


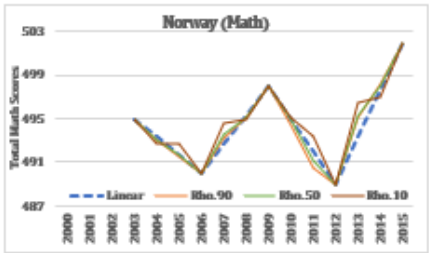


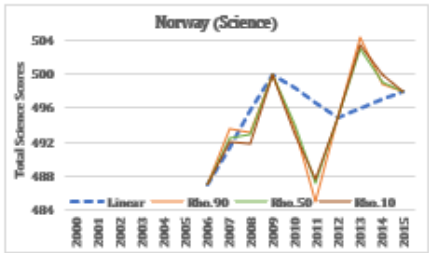


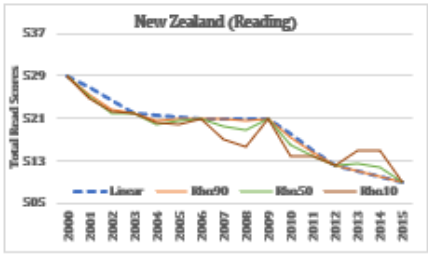


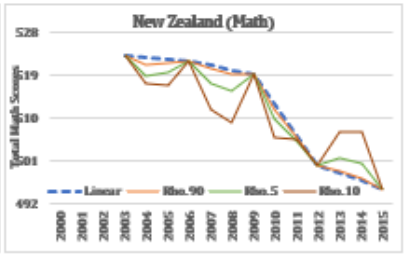


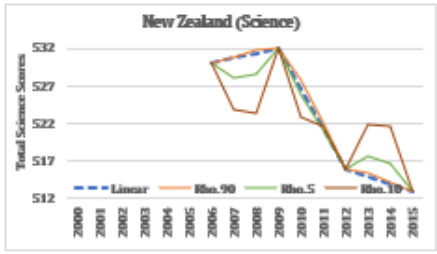




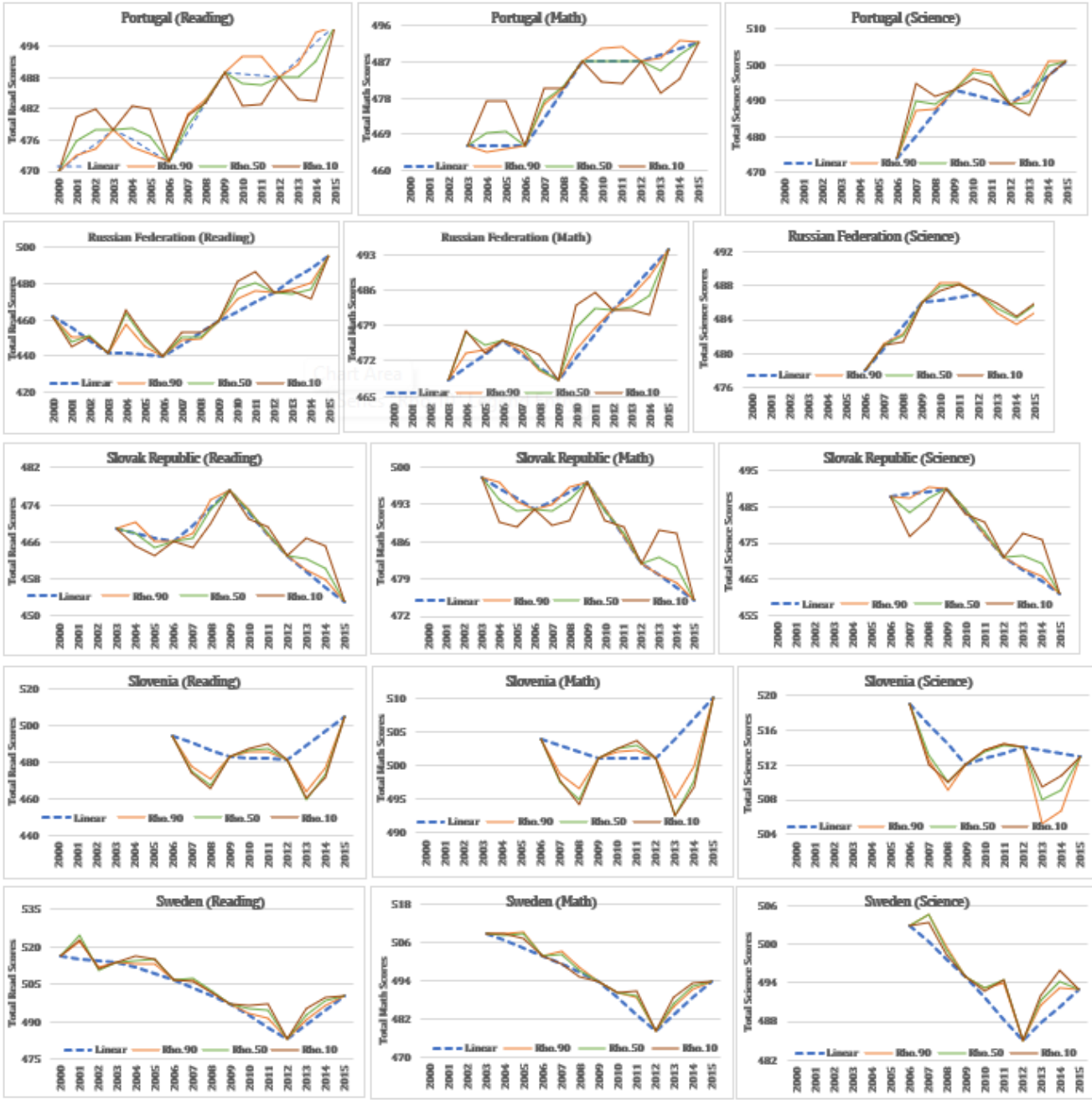


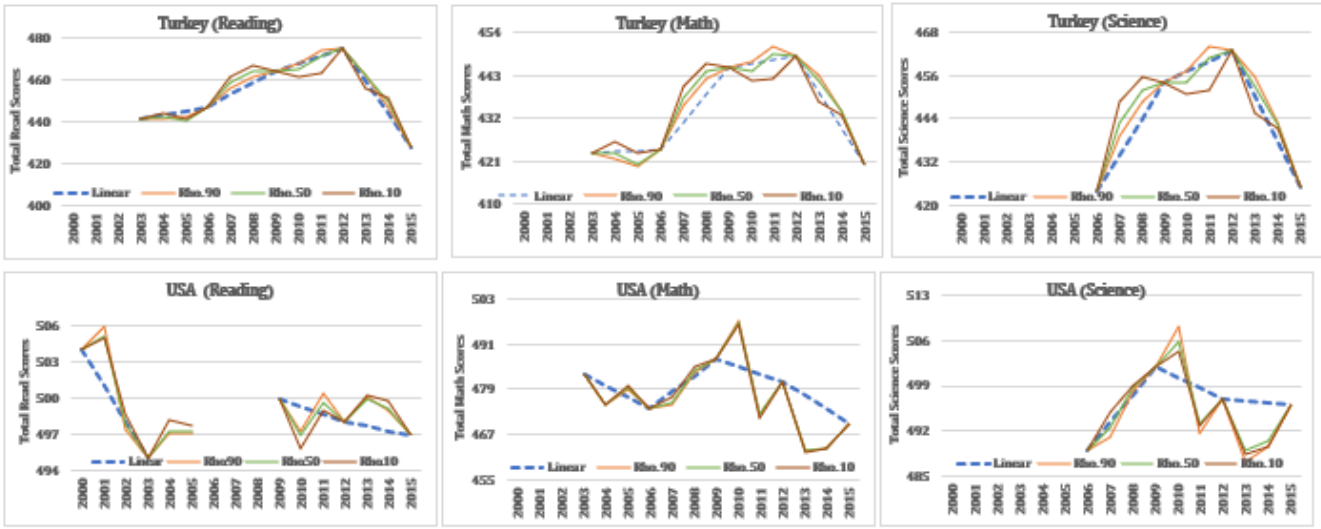






UNIVERSITY of TASMANIA


Portugal (Reading)
Portugal (Math)
Portugal (Science)
Russian Federation (Reading)
Russian Federation (Math)
Russian Federation (Science)
Slovak Republic (Reading)
Slovak Republic (Math)
Slovak Republic (Science)
Slovenia (Reading)
Slovenia (Math)
Slovenia (Science)
Sweden (Reading)
Sweden (Math)
Sweden (Science)
Linear
Rho.90
Rho.50
Rho.10



UNIVERSITY of TASMANIA

Conclusion

- The main contribution of this paper is to extend the PISA data (on education outcomes) which is collected every three years using factor models and the Chow Lin technique to create annual series of 37 OECD countries from 2000 to 2015. The original observations for these countries were 533 observations and we extend it to 1364 observations; where the reading scores were extended from 208 observations to 541 observations; maths scores were extended from 179 observations to 461 observations and science scores were extended from 146 observations to 362 observations. The data set is available from the authors.
  This critical extension of this data set provides benefit for researchers who wants to explore the issues in the field of education, economic growth, income inequality, and human capital development to mention just a few.

Thank You - Questions?