

# Identifying School Climate Variables Associated with Students' Financial Literacy Outcomes

*A Cross-Country Comparison  
Using PISA 2018 Data*

Tony C. A. Tan



Thesis submitted for the degree of  
Master in Assessment, Measurement and Evaluation  
30 credits

Centre for Educational Measurement  
Faculty of Educational Sciences

UNIVERSITY OF OSLO

Spring 2021



# Identifying School Climate Variables Associated with Students' Financial Literacy Outcomes

*A Cross-Country Comparison  
Using PISA 2018 Data*

Tony C. A. Tan

© 2020 Tony C. A. Tan

Identifying School Climate Variables Associated with Students' Financial Literacy Outcomes

<http://www.duo.uio.no/>

Printed: Reprosentralen, University of Oslo

# 敬致父母

To my parents

*Study hard what interests you the most in the  
most undisciplined, irreverent and original manner  
possible.*

*Richard P. Feynman*

# Contents

<b>Contents</b>	<b>i</b>
<b>List of Tables</b>	<b>iii</b>
<b>List of Figures</b>	<b>v</b>
<b>1 Conceptual Framework</b>	<b>1</b>
1.1 In-depth definitions of “financial literacy”	1
1.1.1 Every term my readers need in order to understand my research question	1
1.1.2 Survey not only PISA but also alternative definitions, even critiques of such definitions	1
1.1.3 Any practices that are common in maths/literature but uncommon in financial literacy? Meaning? Implies?	1
1.2 Country-level Financial Knowledge Index	2
<b>2 Methods</b>	<b>5</b>
2.1 Data / Sample / Participants	5
2.2 Measurement of financial literacy	5
2.2.1 Background questions	5
2.2.2 Students’ motivation of spending money	5
2.2.3 Four-point Likert scale	5
2.2.4 Averages	5
2.3 Country-level Financial Knowledge Index	6
2.3.1 Data Collection and Missing Data Treatment	6
2.3.2 Standardisation, Weights and FKI	13
2.4 What exactly I was using to address my research question	15
2.4.1 Sum score? Averages? One particular question?	15
2.4.2 Factor loading? Latent variables?	15
2.4.3 Motivation for choosing these measures	15
2.5 Software and version	15
2.6 My models	15
2.6.1 Motivation for choosing this particular model	15
2.6.2 Refer to my research question	15
2.7 Estimators I obtained	15
2.7.1 Motivation why these estimators rather than others	15
2.8 Weights? Plausible values?	15
2.9 Missing data and how I treated missing data	15
2.10 Model comparison	15
2.11 Guidelines and indices	15
<b>Appendices</b>	<b>19</b>
<b>A GDPR Documentation and Ethical Approval</b>	<b>21</b>
<b>B Analysis Code</b>	<b>25</b>
B.1 Chapter 1	25
B.2 Chapter 2	25
B.2.1 Data Import	25

B.2.2	Missing Pattern Inspection . . . . .	26
B.2.3	Financial Knowledge Index . . . . .	27
B.2.4	Data Reimport . . . . .	33
<b>C</b>	<b>Derivation of Moderated Mediation Effect</b>	<b>37</b>
C.1	Models with Mediators Only . . . . .	37
C.2	Models with Moderated Mediators . . . . .	37
C.3	Mplus Execution . . . . .	39



# List of Tables

1.1	Percentages of Missing Values . . . . .	3
2.1	Data Sources for FKI Computation . . . . .	8
2.2	Data Utilised for Computing FKI . . . . .	12
2.3	FKI and Sub-indices . . . . .	14



# List of Figures

1.1	Three-level Structural Equation Model Predicting Youth's Financial Literacy Outcomes . . . . .	4
2.1	Proportion of Postgraduates to Total Tertiary Graduations . . . . .	10
2.2	Time Series Trend Test . . . . .	11
C.1	Moderated Mediation Model . . . . .	41



# Acknowledgement

Thank-you goes to



# Popular Abstract

This is a press release style abstract.





# Abstract

Repeated economic crises in recent memory have exposed the harsh consequences of financial *illiteracy* shared by high proportions of the general population. Policy makers experienced little resistance when identifying youth as the most effective group for bringing about improvement in citizens' ability to engage with economic and financial matters, but opinions quickly diverge over the optimal approaches for achieving such targeted outcome. Existing literature frequently reports the importance of family environment in cultivating students' financial literacy through the process of "financial socialisation" – [definition goes here] (reference). Such practice, however, encounters interrogation by educators over equity concerns should families remain the main arena for financial literacy development. Schools play vital roles in alleviating inequality in accessing education and training in general but scarce research so far has been devoted into identifying the specific classroom factors that are most effective in advancing students' financial literacy outcomes. The current study therefore attempts to contribute to this enquiry by investigating the relationship between school climate variables and students' financial literacy achievement with an aim of stimulating policy debate about the levers and instruments available to education interventionists for the purpose of improving young people's financial literacy and preparedness as they step into an increasingly uncertain world. Using the 2018 PISA dataset, this paper employs a three-level hierarchical model to conduct cross-country comparisons to highlight school climate variables that are most strongly associated with high financial literacy outcomes.



# Chapter 1 Conceptual Framework

## 1.1 In-depth definitions of “financial literacy”

1.1.1 Every term my readers need in order to understand my research question

1.1.2 Survey not only PISA but also alternative definitions, even critiques of such definitions

1.1.3 Any practices that are common in maths/literature but uncommon in financial literacy? Meaning? Implies?

## 1.2 Country-level Financial Knowledge Index

PISA 2018 financial literacy dataset (OECD, 2020) provides rich information about students and schools. For the purpose of cross-country comparison, however, the country-level financial literacy information must be addressed separately by the researchers. Earlier attempts such as Moreno-Herrero et al. (2018) approximated this information using a variable “quality of math and science education” to control for country-level differences since consensus is yet to emerge about the most appropriate measure for countries’ financial knowledge. Inspired by the UN’s approach to forming Human Development Indices, a recent publication by Oliver-Márquez et al. (2020) proposed a macroeconomic measure for countries’ general financial knowledge levels by examining their economic capability, educational training, existing practices in the financial markets as well as incentives to interact with financial products. More specifically, the authors considered a country’s economic capability, represented by its GDP per capita, to be a key dimension in bringing about its financial knowledge index (FKI). Secondly, literature converges on the importance of educational training for a country’s financial knowledge capability (OECD, 2005). Thirdly, countries with regular engagement with sophisticated financial products and financial markets should possess higher FKI. Lastly, countries with higher aggregate consumption levels and with ageing populations are likely to possess higher FKI due to more frequent exposure and pressure in retirement provision, respectively. Macroeconomic data needed for these computations can be sourced from the World Bank (World Bank, 2020) and the United Nations’ *Human Development Reports* (United Nations, 2020).

Combining individual and institutional data sources can be a productive approach in international large-scale assessment (ILSA) research. According to the framework for comparative education analyses (Bray & Thomas, 1995), this project extends education outcome measures to a country level, addresses the aspect of society and labour market, and relates countries’ entire populations to ILSA research (Strietholt & Scherer, 2018). By combining education outcome data with countries’ economic performance indicators, this project remains most comparable to Hanushek and Woessmann (2012)—while these authors looked into the relationship between countries’ education achievement and their GDP growth, the current investigation highlights how countries’ GDP, along with other macroeconomic practices, in turn systematically impacts on their youth’s educational performance.

**Table 1.1**  
*Percentages of Missing Values*

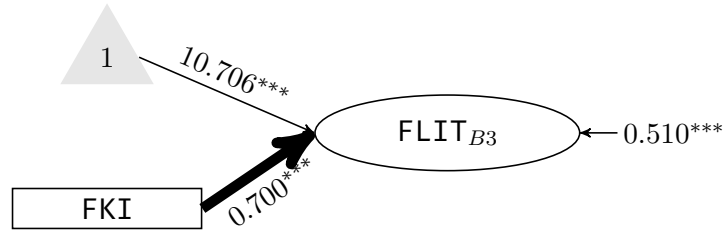
CNT	MALE	IMMI1GEN	IMMI2GEN	ESCS	FCFMLRTY	FLCONFIN	PERFEED	TEACHINT	FLSCHOOL	DISCRIM <sup>†</sup>	BELONG	BULLY	FLFAMILY	CURSUPP <sup>†</sup>	PASCHPOL <sup>†</sup>	STRATIO	EDUSHT	STAFFSHT
BGR	0	6	6	3	12	27	10	10	21	28	19	31	22	100	100	8	3	3
BRA	0	5	5	2	12	34	9	8	21	36	23	40	24	17	19	12	6	7
CAN <sup>†</sup>	0	7	7	5	11	15	100	100	13	100	8	14	14	100	100	100	2	2
CHL	0	4	4	3	10	24	5	4	13	30	15	34	15	9	8	18	9	9
ESP	0	3	3	2	5	21	3	2	7	25	9	29	8	100	100	11	5	6
EST	0	3	3	3	4	8	3	3	6	9	5	11	6	100	100	0	0	0
FIN	0	2	2	2	4	10	3	3	6	100	6	11	7	100	100	2	7	7
GEO	0	5	5	2	9	26	9	9	17	100	15	22	21	4	5	1	2	2
IDN	0	3	3	1	3	6	3	2	5	3	2	5	5	100	100	23	14	14
ITA	0	4	4	3	7	17	4	4	10	23	10	27	12	16	17	9	3	3
LTU	0	3	3	3	4	12	3	3	5	17	8	20	7	100	100	0	0	0
LVA	0	2	2	2	5	9	3	3	6	14	6	15	7	100	100	6	3	4
NLD	0	3	3	2	3	5	3	2	4	100	4	8	4	100	100	11	5	5
PER	0	2	2	1	2	11	5	4	4	56	31	65	5	100	100	2	0	0
POL	0	1	1	1	3	7	2	1	5	9	3	11	5	100	100	0	0	0
PRT	0	6	6	5	8	11	6	6	10	15	8	17	10	10	10	11	1	1
RUS	0	3	3	2	8	13	5	4	11	13	8	15	11	100	100	3	3	3
SRB	0	3	3	1	10	25	8	7	18	25	15	27	19	100	100	8	1	1
SVK	0	2	2	1	4	12	4	3	7	14	6	17	8	100	100	6	6	7
USA	0	3	3	2	3	6	2	1	4	100	4	6	4	100	100	16	10	10

*Note.* Using shades of red in addition to numbers (measured in %), this table visualises the missing percentages by variable and by country. Variables DISCRIM, CURSUPP and PASCHPOL are no longer pursued in the model because too many countries chose not to respond to these questions. Canada (CAN) is not included due to 100 percent missings on multiple variables. <sup>†</sup> marks the country and variables that are excluded from subsequent analyses.

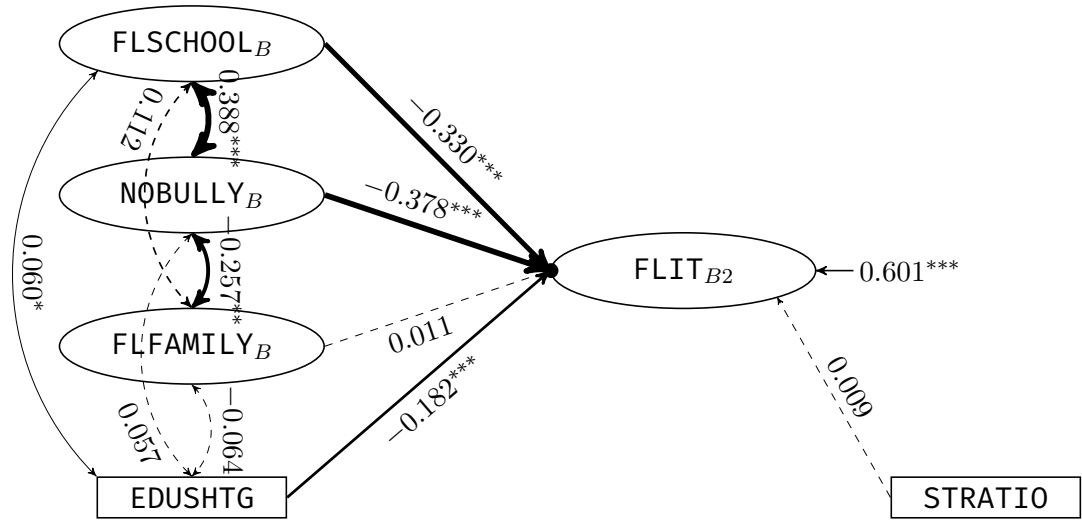
**Figure 1.1**

*Three-level Structural Equation Model Predicting Youth's Financial Literacy Outcomes*

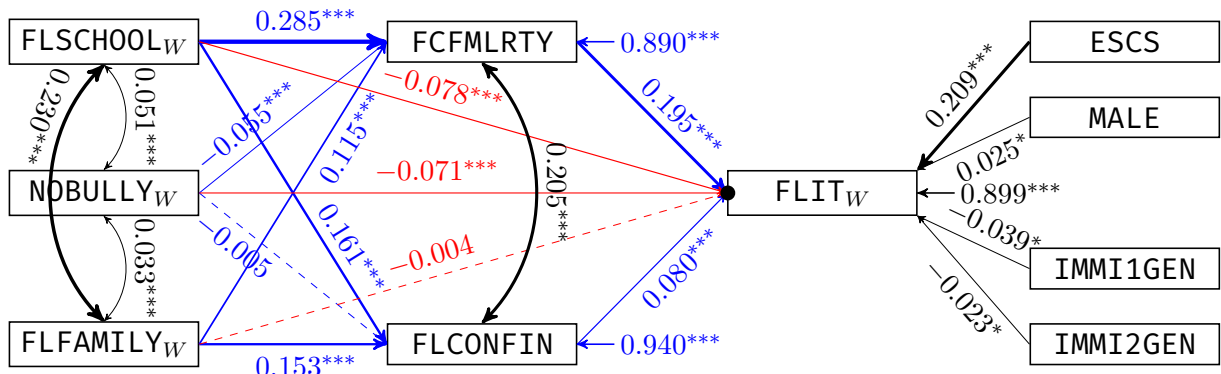
**L3: Country** .....



**L2: School** .....



**L1: Student** .....



*Note.* This multilevel structural equation model predicts youth's financial literacy outcomes using PISA 2018 data, with mediating effects of familiarity with, and confidence in, financial matters at student-level (direct and indirect pathways). Statistics are standardised regression coefficients. Numbers following short arrows  $\leftarrow$  stand for residual variances. Dashed lines represent nonsignificant at  $\alpha = 0.05$  level. FKI = financial knowledge indices, FLIT = financial literacy, subscript  $W$  = within,  $B$  = between.

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

# Chapter 2 Methods

## 2.1 Data / Sample / Participants

This study drew its primary data source from PISA 2018 database (OECD, [2020](#)) containing 107,174 observations spanning 20 countries, in which students were asked about their demographic background, family lives and school experiences. For the financial literacy section, in particular, students responded to questions about their confidence about financial matters, familiarity with concepts of finance, their parental involvement in matters of financial literacy. Ten plausible values were subsequently generated by PISA organisers as measures of students' financial literacy outcomes and were used as the dependent variable.

Student-level independent variables are

School-level independent variables are

Country-level independent variables are

Missing data are handled using Mplus's multiple imputation procedure with ten imputations generated and pooled subsequently following Rubin's Rule (Rubin, 1976).

A three-level multigroup structural equation model was employed to account for the hierarchical structure of the PISA design, with private versus public school as the grouping variable.

## 2.2 Measurement of financial literacy

### 2.2.1 Background questions

### 2.2.2 Students' motivation of spending money

### 2.2.3 Four-point Likert scale

### 2.2.4 Averages

## 2.3 Country-level Financial Knowledge Index

This project closely follows Oliver-Márquez et al.’s (2020) procedure in developing country-level financial knowledge indices using four sub-indices: economic capability (**EC**), educational training (**ET**), existing practices in financial market (**Use**), and incentives (**Need**) to engage with financial products. The first sub-index **EC** is calculated using the logarithm of a country’s GDP per capita in current international dollars (purchasing power parity adjusted). For the **ET** sub-index, a country’s highly skilled workforce is represented by its postgraduate to total tertiary graduation ratio, and the mean years of schooling is used to measure its general education level. For the **Use** sub-index, gross portfolio equity assets (GPEA) and insurance company assets (ICA) are considered sophisticated financial products a country engages in. Additionally, in order to capture the central role of technology in amplifying the proliferation and use of financial assets, the proportion of a country’s Internet users (IUS) enters the definition via

$$\text{Use} = (\text{GPEA} + \text{ICA})^{\text{IUS}}.$$

The final sub-index **Need** is compiled as

$$\text{Need} = (\text{PFA} + \text{AC})^{\text{AGEING}},$$

where PFA is the pension fund assets to GDP ratio. Aggregate consumption is defined as:

$$\text{AC} = \frac{2\% \times \text{household final consumption expenditure}}{\text{GDP}},$$

with the “2% rule” being drawn from Caliendo and Findley’s (2013) derivation, and the proportion of ageing population is computed as

$$\text{AGEING} = \frac{\left[ \frac{\text{population}(>65)}{\text{population}(20\sim64)} \right]_{2018} - \left[ \frac{\text{population}(>65)}{\text{population}(20\sim64)} \right]_{2009}}{\left[ \frac{\text{population}(>65)}{\text{population}(20\sim64)} \right]_{2009}}.$$

### 2.3.1 Data Collection and Missing Data Treatment

The data sources for FKI computation are documented in Table 2.1 and its associated notes.

Sub-indices **ET** and **Use** both contain missing observations for the year 2018. Majority of



such missing data appear to be the result of administrative delay, with historic observations available until 2017. It is therefore feasible to conduct time-series forecasts using prior year observations to best approximate 2018 values.

**Table 2.1***Data Sources for FKI Computation*

Database <sup>a</sup>	Country <sup>b</sup>	Series	Time
Economic Capacity			
WB-dev	20	GDP per capita, PPP (current international \$)	2018
Educational Training			
WB-ed	20 \ Russia	Graduates from ISCED 7 programmes in tertiary education, both sexes (number)	2013– <b>2018</b>
		Graduates from ISCED 8 programmes in tertiary education, both sexes (number)	2013– <b>2018</b>
		Graduates from tertiary education, both sexes (number)	2013– <b>2018</b>
RS	Russia	PhD (Type 1) <sup>c</sup> , PhD (Type 2) <sup>d</sup>	2018
RE	Russia	Master (Type 1) <sup>e</sup> , Master (Type 2) <sup>f</sup> , total tertiary <i>excluding</i> PhD <sup>g</sup>	2018
HDR	20	Dimension = Education; Education = Mean years of schooling (years)	2018
Use			
WB-fin	20	Gross portfolio equity assets to GDP (%)	2011– <b>2018</b>
		Insurance company assets to GDP (%)	2011– <b>2018</b>
WB-dev	20	Individuals using the Internet (% of population)	2009– <b>2018</b>
Need			
WB-fin	20 \ Georgia	Pension fund assets to GDP (%)	2008– <b>2018</b>
GP	Georgia	Minutes of the meeting of the investment board of the Pension Agency <sup>h</sup>	2019 <sup>*</sup>
GS	Georgia	GDP at current prices, billion GEL <sup>i</sup>	2018
WB-dev	20	Household and NPISHs final consumption expenditure, PPP (current international \$)	2018
		GDP, PPP (current international \$)	2018
		Population ages 0–14, male	2009, 2018
		Population ages 0–14, female	2009, 2018
		Population ages 15–64, male	2009, 2018
		Population ages 15–64, female	2009, 2018
		Population ages 65 and above, male	2009, 2018
		Population ages 65 and above, female	2009, 2018
		Population ages 15–19, male (% of male population)	2009, 2018
		Population ages 15–19, female (% of female population)	2009, 2018

*Note.* Sub-indices are shaded in gray. Bold font signifies this year contains missing data.

- <sup>a</sup> WB-dev = [World Bank – World development indicators](#)  
 WB-ed = [World Bank – Education statistics – All indicators](#)  
 WB-fin = [World Bank – Global financial development](#)  
 HDR = [Human Development Reports – Data](#)  
 RS = [Russian Federal State Statistic Service](#)  
 RE = [Russian Ministry of Education and Science](#)  
 GP = [Pension Agency of Georgia](#)  
 GS = [National Statistics Office of Georgia](#)
- <sup>b</sup> “20” = the 20 participating countries in 2018 PISA financial literacy test: Brazil, Bulgaria, Canada, Chile, Estonia, Finland, Georgia, Indonesia, Italy, Latvia, Lithuania, the Netherlands, Peru, Poland, Portugal, Russian Federation, Serbia, Slovak Republic, Spain, and the USA. “\” = excluding or except
- <sup>c</sup> [https://rosstat.gov.ru/storage/mediabank/asp-2\(1\).xls](https://rosstat.gov.ru/storage/mediabank/asp-2(1).xls), Sheet “по направлениям подготовки”, Cell C7 = number of PhD graduates (Type 1)
- <sup>d</sup> <https://rosstat.gov.ru/storage/mediabank/asp-3.xls>, Sheet “по научным специальностям”, Cell B7 = number of PhD graduates (Type 2)
- <sup>e-g</sup> [https://minobrnauki.gov.ru/common/upload/download/VPO\\_1\\_2018.rar](https://minobrnauki.gov.ru/common/upload/download/VPO_1_2018.rar) contains a spreadsheet [СВОД\\_БПО1\\_БСЕГО.xls](#), Sheet “P2\_1\_3(1)”, Cell E198 = number of master graduates (Type 1)<sup>e</sup>, Cell E410 = number of master graduates (Type 2)<sup>f</sup>, Cell E592 = total tertiary graduates *excluding* PhD<sup>g</sup>
- <sup>h</sup> [Minutes of the meeting of the investment board of the Pension Agency](#), p. 4, no. 3
- <sup>i</sup> [Gross domestic product \(GDP\)](#), row = GDP at current prices, billion GEL, column = 2018
- <sup>\*</sup> Georgia started a [new pension system](#) on 1 January 2019. Since 2018 was a transitional period with scarce data, 2019 is used as the best approximation for Georgia’s pension system for 2018.

## Sub-index ET

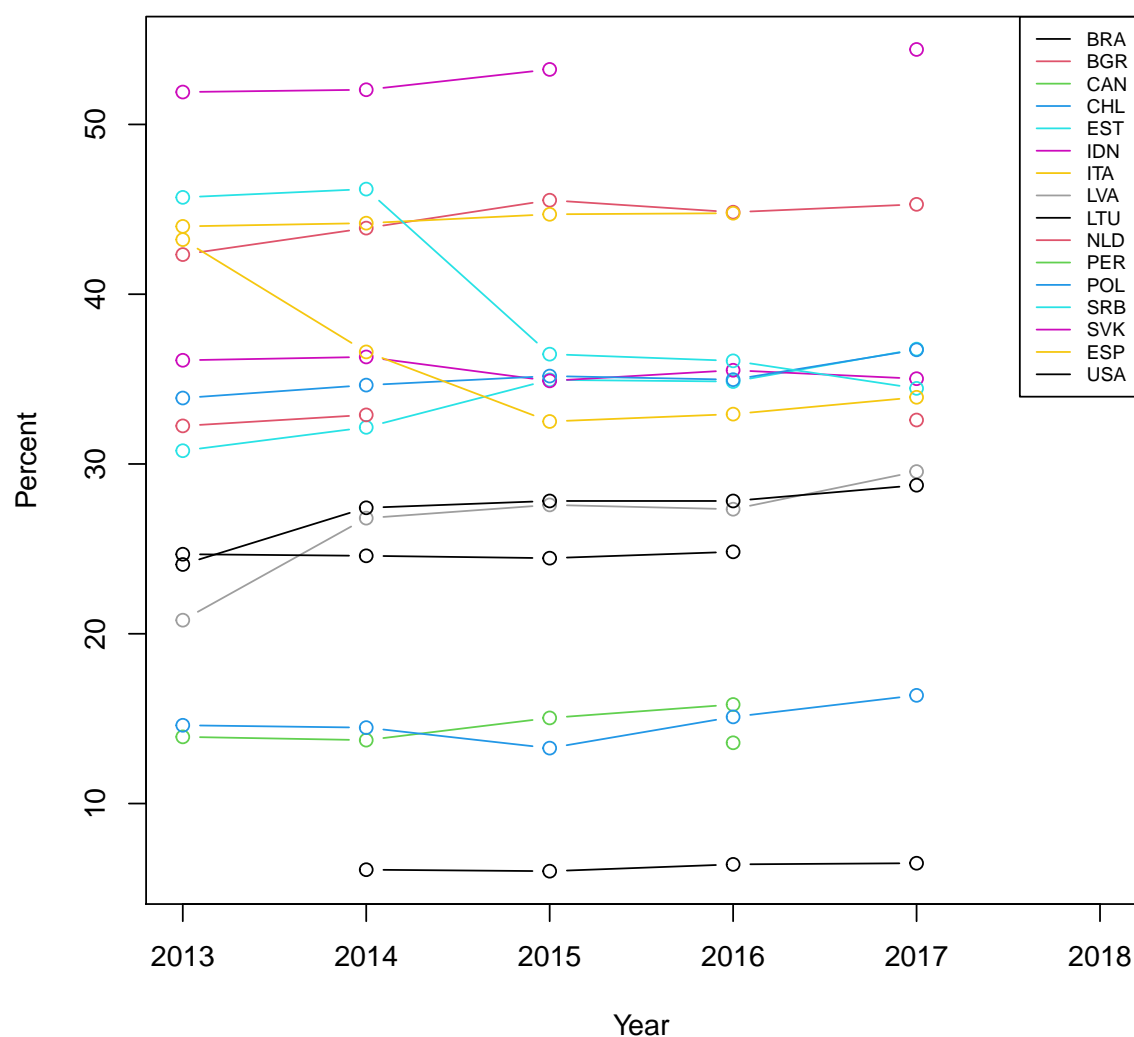
The 2018 archive for the number of master (ISCED 7), PhD (ISCED 8), and total tertiary graduates are incomplete for all participating countries except Georgia, Indonesia and Serbia.

Figure 2.1 presents a time series plot of

$$\text{SKILLED} = \frac{\text{number of masters} + \text{number of PhDs}}{\text{total number of tertiary graduates}}$$

**Figure 2.1**

*Proportion of Postgraduates to Total Tertiary Graduations*



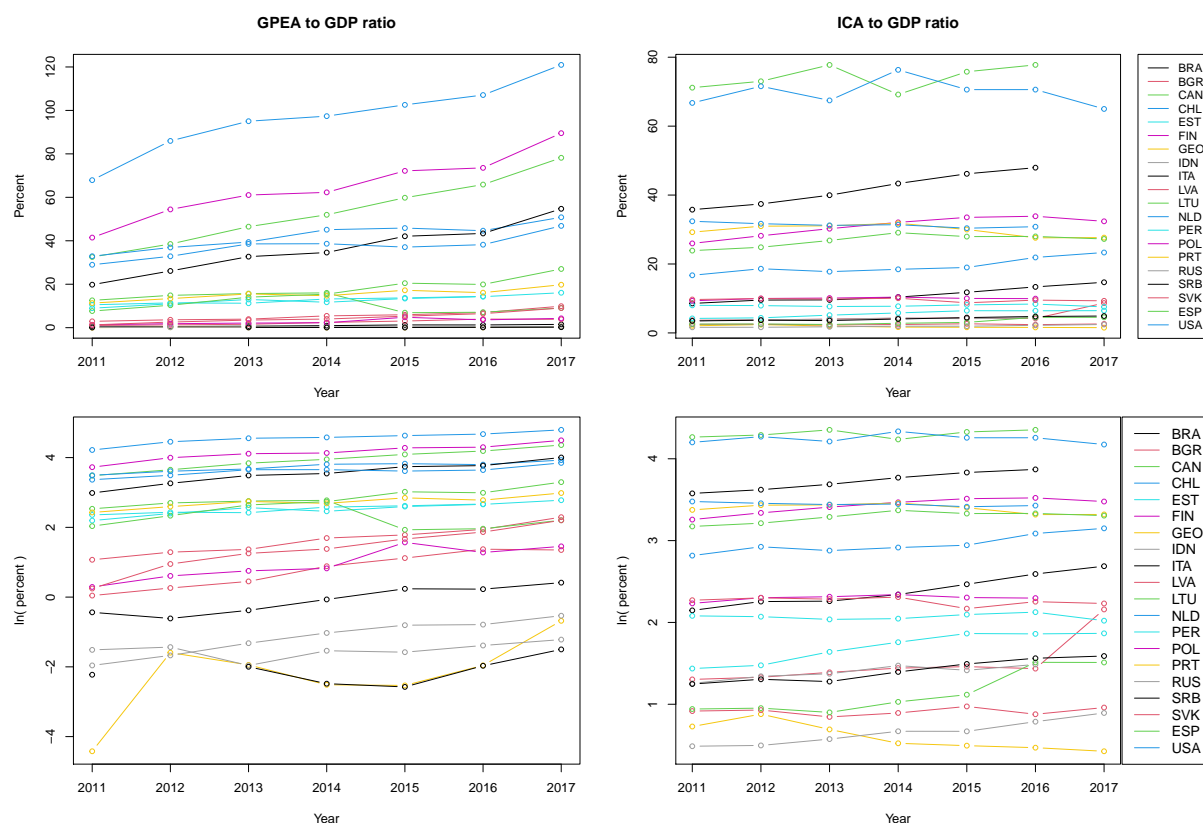
*Note.* “Postgraduate” is defined as master (ISCED 7) and PhD (ISCED 8) graduates. Countries not shown: GEO, IDN and SRB (2018 data available) and RUS (consult other sources)

and suggests that this ratio is likely to be stable over time, especially between adjacent years. A “naive forecast”, where the nearest available year’s data are to be duplicated for 2018, is applied for SKILLED.

## Sub-index Use

All series involved in calculating this sub-index, GPEA, ICA and IUS, contain missing data. When time series data contain only exponential growth but no underlying trend, a simple exponential smoothing would suffice (Gardner, 1985); if trend is present, Holt-Winters method is superior (Chatfield, 1978). Figure 2.2 facilitates this decision making by plotting both the original and log-transformed versions of GPEA and ICA series. Since curves after log-transformations have slopes, it is prudent to apply the Holt-Winters forecasting method in order to account for possible trends contained in the original series.

**Figure 2.2**  
*Time Series Trend Test*



*Note.* The time series plots after natural logarithm transformations (bottom panels) are not flat, suggesting the original series (top panels) contain trends. Holt-Winters method therefore is preferred over simple exponential smoothing for 2018 forecasts.

The IUS series contains missing data for Canada, Chile and the United States. Similar Holt-Winters procedure is applied to recover 2018 IUS data.

**Table 2.2***Data Utilised for Computing FKI*

	Economic Capacity	Educational Training		Use			Need		
	GDP per capita	Skilled	Schooling	GPEA	ICA	IUS	PFA	AC	AGEING
BRA	9.612	6.484	7.8	1.683	16.259	70.434	11.827	1.210	0.288
BGR	10.026	45.294	11.8	4.114	7.044	64.782	13.577	1.091	0.234
CAN	10.821	15.832	13.3	84.010	77.728	93.588	96.205	1.068	0.271
CHL	10.117	16.371	10.4	51.755	25.591	89.531	73.225	1.073	0.214
EST	10.501	36.765	13	16.399	7.681	89.357	18.012	0.876	0.163
FIN	10.807	35.024	12.4	93.626	31.481	88.890	52.024	0.974	0.370
GEO	9.588	24.039	12.8	0.784	1.469	62.718	0.834	1.227	0.042
IDN	9.362	7.771	8	0.636	4.612	39.905	1.826	1.059	0.145
ITA	10.665	44.771	10.2	57.434	51.260	74.387	10.589	1.075	0.155
LVA	10.330	29.554	12.8	8.598	2.538	83.577	14.732	1.027	0.142
LTU	10.487	28.749	13	9.008	5.500	79.723	7.457	1.107	0.149
NLD	10.961	32.590	12.2	124.171	64.956	94.712	207.938	0.805	0.326
PER	9.479	13.577	9.2	16.027	6.505	52.540	22.530	1.187	0.227
POL	10.368	36.725	12.3	4.853	9.535	77.542	9.838	1.085	0.355
PRT	10.444	34.454	9.2	19.353	25.579	74.661	8.761	1.133	0.237
RUS	10.267	30.349	12	0.302	2.614	80.865	4.415	0.941	0.155
SRB	9.774	26.946	11.2	0.306	5.111	73.361	0.845	1.171	0.280
SVK	10.391	54.417	12.6	10.644	8.873	80.660	12.497	0.962	0.300
ESP	10.609	33.929	9.8	27.681	28.230	86.107	10.235	1.044	0.186
USA	11.048	24.825	13.4	55.505	30.183	84.881	150.040	1.364	0.252

*Note.* Full variable names: Skilled = Postgraduate to total tertiary ratio; Schooling = Mean year of schooling; GPEA = Gross portfolio to GDP ratio; ICA = Insurance company assets to GDP ratio; IUS = Number of Internet users per 100 population; PFA = Pension fund assets to GDP ratio; AC = 2% of household final consumption expenditure to GDP ratio; AGEING = Aged-to-productive-population ratio (% change between 2009 and 2018)

## Other Items with Data Concerns

Russia reported 67.96% and 61.01% of its total university degree recipients to be postgraduates for the year 2013 and 2015 respectively (2014 missing). This figure rapidly declines to 41.6% in 2016 and further down to 25.69% in 2017. Such volatility goes against the stable patterns shared by most countries in [Figure 2.1](#), casting doubt on data reliability. Separate investigation is therefore conducted using Russian government archive (Notes c to g in [Table 2.1](#)).

Georgia underwent pension reform in 2018 with fund balance gradually transitioning to State Pension Agency for its official resumption of duty on 1 January 2019. Resultantly, 2018 pension balance for this country is unavailable but to be best approximated using 2019 official data (Notes h, i and \* of [Table 2.1](#)).

[Table 2.2](#) documents the results of the abovementioned data recovery process.

### 2.3.2 Standardisation, Weights and FKI

Following Oliver-Márquez et al. ([2020](#))’s procedure, all series in [Table 2.2](#) undergo min-max normalisation such that the smallest entry receives a new score of 0.01 and the biggest number is re-coded to 0.99. This slight deviation from the original paper (where the min-max normalisation yields 0 to 1) is to avoid multiplying a series by zero or raising a base to the power of zero.

Variable weights are calculated following Oliver-Márquez et al. ([2020](#))’s recipe to be the inverses of each series’ standard deviations. Whereas a sub-index combines more than one series, each weight is further divided by the sum of the constituent weights so that total weights add to one.

FKI is finally computed by taking the geometric mean of all four sub-indices, subject to sub-index-weights similar to variable weights above, as presented in [Table 2.3](#).

**Table 2.3***FKI and Sub-indices*

	FKI	EC	ET	Use	Need
NLD	0.940	0.939	0.640	1.805	1.000
USA	0.937	0.990	0.589	0.856	1.406
CAN	0.784	0.858	0.409	1.637	0.953
ITA	0.762	0.767	0.602	1.069	0.807
FIN	0.724	0.850	0.685	1.127	0.562
ESP	0.627	0.735	0.464	0.635	0.726
LTU	0.613	0.664	0.632	0.243	0.836
PRT	0.591	0.639	0.401	0.630	0.762
BGR	0.583	0.396	0.760	0.384	0.729
EST	0.577	0.672	0.746	0.266	0.575
SVK	0.559	0.608	0.924	0.301	0.441
POL	0.555	0.595	0.699	0.286	0.572
LVA	0.550	0.573	0.633	0.161	0.795
CHL	0.544	0.449	0.302	0.761	0.908
RUS	0.450	0.536	0.597	0.083	0.639
GEO	0.424	0.141	0.547	0.210	0.997
SRB	0.423	0.249	0.500	0.193	0.742
PER	0.309	0.078	0.194	0.691	0.877
BRA	0.141	0.155	0.010	0.432	0.833
IDN	0.122	0.010	0.040	0.973	0.787

*Note.* Table sorted in descending order by countries'  
FKI. FKI = financial knowledge index, EC =  
Economic Capability, ET = Educational Training.



## **2.4 What exactly I was using to address my research question**

### **2.4.1 Sum score? Averages? One particular question?**

### **2.4.2 Factor loading? Latent variables?**

### **2.4.3 Motivation for choosing these measures**

## **2.5 Software and version**

## **2.6 My models**

### **2.6.1 Motivation for choosing this particular model**

### **2.6.2 Refer to my research question**

## **2.7 Estimators I obtained**

### **2.7.1 Motivation why these estimators rather than others**

## **2.8 Weights? Plausible values?**

## **2.9 Missing data and how I treated missing data**

## **2.10 Model comparison**

## **2.11 Guidelines and indices**



# References

- Bray, M. & Thomas, R. M. (1995). Levels of comparison in educational studies: Different insights from different literatures and the value of multilevel analyses. *Harvard Educational Review*, 65(3), 472–491. <https://doi.org/10.17763/haer.65.3.g3228437224v4877>
- Caliendo, F. N. & Findley, T. S. (2013). Time inconsistency and retirement planning. *Economics Letters*, 121(1), 30–34. <https://doi.org/10.1016/j.econlet.2013.06.041>
- Chatfield, C. (1978). The Holt-Winters forecasting procedure. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 27(3), 264–279. <https://www.jstor.org/stable/pdf/2347162.pdf>
- Gardner, E., Jr. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, 4, 1–28. <https://www.bauer.uh.edu/gardner/Exp-Sm-1985.pdf>
- Hanushek, E. A. & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267–321. <https://doi.org/10.1007/s10887-012-9081-x>
- Moreno-Herrero, D., Salas-Velasco, M. & Sánchez-Campillo, J. (2018). Factors that influence the level of financial literacy among young people: The role of parental engagement and students' experiences with money matters. *Children and Youth Services Review*, 95, 334–351. <https://doi.org/10.1016/j.childyouth.2018.10.042>
- OECD. (2005). *Improving financial literacy: Analysis of issues and policies*. OECD Publishing. <https://www.oecd-ilibrary.org/docserver/9789264012578-en.pdf>
- OECD. (2020). *Financial literacy data file* (Data set). OECD Publishing. [https://webfs.oecd.org/pisa2018/SPSS\\_STU\\_FLT.zip](https://webfs.oecd.org/pisa2018/SPSS_STU_FLT.zip)
- Oliver-Márquez, F. J., Guarnido-Rueda, A. & Amate-Fortes, I. (2020). Measuring financial knowledge: A macroeconomic perspective. *International Economics and Economic Policy*, 1–46. <https://doi.org/10.1007/s10368-020-00482-2>
- Strietholt, R. & Scherer, R. (2018). The contribution of international large-scale assessments to educational research: Combining individual and institutional data sources. *Scandinavian Journal of Educational Research*, 62(3), 368–385. <https://doi.org/10.1080/00313831.2016.1258729>

United Nations. (2020). *Human development reports* (Data set). The United Nations. <http://hdr.undp.org/en/data>

World Bank. (2020). *World Bank DataBank* (Data set). The World Bank. <https://databank.worldbank.org/home.aspx>

# Appendices



# Appendix A GDPR Documentation and Ethical Approval

This research project discharges its duty imposed by the EEA's general data protection regulation (GDPR) by following Norwegian Centre for Research Data (NSD)'s [notification test](#) on Friday, 11 September 2020. Both [PISA 2018 Database](#) and the [World Bank Open Data](#) contain only aggregated and de-personalised datasets with no possibility of back-tracing to any particular participant. Resultantly, no identifiable personal data were collected or used at any stage of this research. The NSD's assessment letter outlines the agency's decision of not subjecting this project to the GDPR notification. The NSD decision letter also satisfies University of Oslo's [ethical approval requirement](#) and concludes the approval process.

About us ([/personvernombud/en/about\\_us.html](/personvernombud/en/about_us.html))

Norwegian ([/personvernombud/meld\\_prosjekt/meldeplikttest.html](/personvernombud/meld_prosjekt/meldeplikttest.html))

NSD (</>) > Personvern tjenester (</personvernombud/>) > Data Protection Services (</personvernombud/en/>) > Notify project (</personvernombud/en/notify/>) > Notification Test

Denne siden på norsk ([/personvernombud/meld\\_prosjekt/meldeplikttest.html](/personvernombud/meld_prosjekt/meldeplikttest.html))

## Will you be processing personal data?

Are you unsure whether your project is subject to notification? Feel free to try our informal Notification test. Note that the test is intended as a guidance and is not a formal assessment.

**Will you be collecting/processing directly identifiable personal data?** ☐ Yes ☒ No

A person will be directly identifiable through name, social security number, or other uniquely personal characteristics.

Read more about personal data (</personvernombud/en/help/vocabulary.html?id=8>) and notification (</personvernombud/en/notify/index.html>).

NB! Even though information is to be anonymized in the final thesis/report, check the box if identifying personal data is to be collected/processed in connection with the project.

**Will directly identifiable personal information be linked to the data (e.g. through a reference number which refers to a separate list of names/scrambling key)?** ☐ Yes ☒ No

Note that the project will be subject to notification even if you cannot access the scrambling key (</personvernombud/en/help/vocabulary.html?id=11>), as the procedure often is when using a data processor (</personvernombud/en/help/vocabulary.html?id=3>), or in register-based studies ([/personvernombud/en/help/research\\_methods/register\\_studies.html](/personvernombud/en/help/research_methods/register_studies.html)).

**Will you be collecting/processing background information that may identify individuals (indirectly identifiable personal data)?** ☐ Yes ☒ No

A person will be indirectly identifiable if it is possible to identify him/her through a combination of background information (such as place of residence or workplace/school, combined with information such as age, gender, occupation, etc.).

**Will there be registered personal data (directly/indirectly/via IP or email address, etc.) using online surveys?** ☐ Yes ☒ No

Please note that the project will be subject to notification even if you as a student/researcher cannot access the link to the IP or email address, as the procedure often is when using a data processor.



Read more about online surveys ([/personvernombud/en/help/research\\_methods/online\\_surveys.html](/personvernombud/en/help/research_methods/online_surveys.html)).

**Will there be registered personal data using digital photo or video files?**

☐ Yes☒ No

Photo/video recordings of faces will be regarded as identifiable personal data. In order for a voice to be considered as identifiable, it must be registered in combination with other background information, in such a way that people can be recognized.

Show results

## Notify project

Do I have to notify my project? (</personvernombud/en/notify/index.html>)

Notification Form ([/personvernombud/en/notify/meldeskjema\\_link](/personvernombud/en/notify/meldeskjema_link))

Notifying changes ([/personvernombud/en/notify/notifying\\_changes.html](/personvernombud/en/notify/notifying_changes.html))

## Get help notifying your project

Processing the notification (</personvernombud/en/help/index.html>)

Frequently asked questions (</personvernombud/en/help/faq.html>)

Vocabulary (</personvernombud/en/help/vocabulary.html>)

Research topics ([/personvernombud/en/help/research\\_topics/](/personvernombud/en/help/research_topics/))

Research methods ([/personvernombud/en/help/research\\_methods/](/personvernombud/en/help/research_methods/))

Information and consent ([/personvernombud/en/help/information\\_consent/](/personvernombud/en/help/information_consent/))

Other approvals ([/personvernombud/en/help/other\\_approvals/](/personvernombud/en/help/other_approvals/))

## Result of Notification Test: Not Subject to Notification

You have indicated that neither directly or indirectly identifiable personal data will be registered in the project.

If no personal data is to be registered, the project will not be subject to notification, and you will not have to submit a notification form.

Please note that this is a guidance based on information that you have given in the notification test and not a formal confirmation.

For your information: *In order for a project not to be subject to notification, we presuppose that all information processed using electronic equipment in the project remains anonymous.*

*Anonymous information is defined as information that cannot identify individuals in the data set in any of the following ways:*

- directly, through uniquely identifiable characteristic (such as name, social security number, email address, etc.)*
- indirectly, through a combination of background variables (such as residence/institution, gender, age, etc.)*
- through a list of names referring to an encryption formula or code, or*
- through recognizable faces on photographs or video recordings.*

*Furthermore, we presuppose that names/consent forms are not linked to sensitive personal data.*

Kind regards,  
NSD Data Protection

# Appendix B Analysis Code

## B.1 Chapter 1

There is no analysis code in ??.

## B.2 Chapter 2

### B.2.1 Data Import

```
1 | # Housekeeping
2 | library(Orcs)
3 | setwdOS(lin = "~/", win = Sys.getenv("USERPROFILE"))
4 |
5 | # Import SPSS file into R
6 | library(intsvy)
7 | finlit <- pisa.select.merge(
8 |   student.file = "CY07_MSU_FLT_QQQ.SAV", # file ext in capital
9 |   school.file = "CY07_MSU_SCH_QQQ.sav", # file ext in lower case
10 |   student = c(
11 |     # Control variables
12 |     "ST004D01T", # Student (Standardized) Gender
13 |     "IMMIG", # Index Immigration status
14 |     "ESCS", # Index of economic, social and cultural status
15 |     # Mediators
16 |     "FCFMLRTY", # Familiarity with concepts of finance (Sum)
17 |     "FLCONFIN", # Confidence about financial matters (WLE)
18 |     # Academic
19 |     "PERFEED", # Perceived feedback (WLE)
20 |     "TEACHINT", # Perceived teacher's interest (WLE)
21 |     "FLSCHOOL", # Financial education in school lessons (WLE)
22 |     # Safety
23 |     "DISCRIM", # Discriminating school climate (WLE)
24 |     "BELONG", # Sense of belonging to school (WLE)
25 |     "BEINGBULLIED", # Student's experience of being bullied (WLE)
26 |     # Community
27 |     "FLFAMILY", # Parental involvement in matters of Financial Literacy (WLE)
28 |     "CURSUPP", # Current parental support for learning at home (WLE)
29 |     "PASCHPOL" # School policies for parental involvement (WLE)
30 |   ),
31 |   school = c(
32 |     "STRATIO", # Student-teacher ratio
33 |     "EDUSHORT", # Shortage of educational material (WLE)
34 |     "STAFFSHORT" # Shortage of educational staff (WLE)
35 |   ),
36 |   countries = c(
37 |     "BGR", "BRA", "CAN", "CHL", "ESP",
38 |     "EST", "FIN", "GEO", "IDN", "ITA",
39 |     "LTU", "LVA", "NLD", "PER", "POL",
40 |     "PRT", "QMR", "QRT", "RUS", "SRB",
41 |     "SVK", "USA"
42 |   )
43 | )
```

```

44 |
45 | # Inspect table header
46 | names(finlit)
47 |
48 | # Remove columns that I do not need
49 | finlit <- finlit[, -c(5, 7:86)] # 5 = BOOKID; 7:86 = resampling weights
50 |
51 | # Some var need recording
52 | library(car)
53 |
54 | # Re-code Russian territories to RUS
55 | finlit$CNT <- recode(finlit$CNT, "
56 |   'QMR' = 'RUS';
57 |   'QRT' = 'RUS'
58 | ")
59 |
60 | # Recode ST004D01T from Sex to Male
61 | MALE <- finlit$ST004D01T - 1
62 |
63 | # Recode IMMIG to 1st and 2nd generation
64 | IMMI1GEN <- recode(finlit$IMMIG, "
65 |   1 = 0;
66 |   2 = 0;
67 |   3 = 1
68 | ")
69 |
70 | IMMI2GEN <- recode(finlit$IMMIG, "
71 |   1 = 0;
72 |   2 = 1;
73 |   3 = 0
74 | ")
75 |
76 | # Revert coding direction: bigger number => safer school
77 | NOBULLY <- finlit$BEINGBULLIED * (-1)
78 |
79 | # Stitch spreadsheets together
80 | finlit <- cbind(
81 |   finlit[, c(1:35)],
82 |   MALE, IMMI1GEN, IMMI2GEN,
83 |   finlit[, c(38:45)],
84 |   NOBULLY,
85 |   finlit[, c(47:53)]
86 | )
87 |
88 | # Use data.table for better RAM management
89 | library(data.table); setDTthreads(0) # 0 means all the available cores
90 | # Export data into a CSV file for faster import next time
91 | fwrite(finlit,
92 |   file = "finlit.csv",
93 |   na = "NA", row.names = F, col.names = T
94 | )

```

## B.2.2 Missing Pattern Inspection

```

1 | library(Orcs)
2 | setwdOS(lin = "~/", win = Sys.getenv("USERPROFILE"))
3 |
4 | library(data.table); setDTthreads(0)
5 | finlit <- fread("finlit.csv", nThread = getDTthreads())
6 |
7 | library(dplyr)

```

```

8 | # Record how many missings each country has for each var
9 | missings <- finlit %>%
10 |   select(everything()) %>%
11 |   group_by(CNT) %>%
12 |   summarise_all(funs(sum(is.na(.))))
13 | # Give me the headcount for each country
14 | headcount <- finlit %>%
15 |   group_by(CNT) %>%
16 |   summarize(n())
17 | # Stitch these two tables together
18 | missing_table <- tibble(headcount, missings[, -1])
19 | # Save this file
20 | fwrite(missing_table, "missing_table.csv", row.names = F, col.names = T)
21 |
22 | # Inspect the missing table using Excel
23 | # Throw away the following countries
24 | #   CAN: 100% missing on too many var
25 | #   BRA, FIN, LVA, NLD, RUS, SRB: private/public info missing
26 | # Throw away these var
27 | #   DISCRIM, CURSUPP, PASCHPOL: Too many countries have 100% missing

```

## B.2.3 Financial Knowledge Index

```

1 | #Section 0: Housekeeping
2 | library(Orcs) # Set working directory depending on operating system
3 | setwdOS(
4 |   lin = "~/uio/", win = "M:/",
5 |   ext = "pc/Dokumenter/MSc/Thesis/Data/L3/"
6 | )
7 |
8 | # Set up a "bookshelf" to hold variables nessary to compute FKI
9 | fki_raw <- matrix(NA,
10 |   nrow = 20, ncol = 10, dimnames = list(
11 |     c( # row names
12 |       "BRA", "BGR", "CAN", "CHL", "EST",
13 |       "FIN", "GEO", "IDN", "ITA", "LVA",
14 |       "LTU", "NLD", "PER", "POL", "PRT",
15 |       "RUS", "SRB", "SVK", "ESP", "USA"
16 |     ),
17 |     c( # column names
18 |       "gdp_per_capita", # economic capability (sub_ind_ec)
19 |       "highly_skilled", "mean_year_of_schooling", # (sub_ind_et)
20 |       "gpea", "ica", "ius", # use (sub_ind_use)
21 |       "pfa", "ac", "gdp", "ageing" # need (sub_ind_need)
22 |     )
23 |   ) # End list()
24 | ) # End matrix()
25 |
26 |
27 | # Section 1: Economic Capacity (EC)
28 |
29 | gdp_per_capita <- read.csv("gdp_per_capita.csv", header = T, sep = "\t")
30 |
31 | fki_raw[, 1] <- log(gdp_per_capita[, 2])
32 |
33 | rm(gdp_per_capita)
34 |
35 |
36 | # Section 2: Educational Training (ET)
37 |   # Subsection 2.1: Highly skilled
38 |     # Masters

```

```

39| isced_7 ← read.csv("isced_7.csv", header = T, sep = "\t")
40|   # PhDs
41| isced_8 ← read.csv("isced_8.csv", header = T, sep = "\t")
42|   # Total tertiary
43| total_tertiary ← read.csv("total_tertiary.csv", header = T, sep = "\t")
44|
45| # Compute highly skilled (master + PhD) to total tertiary ratio
46| highly_skilled ← ts(
47|   (isced_7 + isced_8) / total_tertiary,
48|   start = 2013, end = 2018, frequency = 1
49| )
50|
51| # # Visualise highly_skilled. Turn off GEO (#7), IDN (#8), SRB (#17) and RUS (#16)
52| # pdf("../Figures/skilled.pdf")
53| # ts.plot(100 * highly_skilled[, -c(7, 8, 16, 17)],
54| #   type = "b", col = 1:15,
55| #   xlab = "Year", ylab = "Percent"
56| # )
57| # legend("topright", colnames(highly_skilled[, -c(6, 7, 15, 16)]),
58| #   col = 1:15, lty = 1, cex = 0.65
59| # )
60| # dev.off()
61|
62| # Decision: naive forecasts, i.e., copy-paste nearest available year
63| library(forecast)
64| # Create a placeholder matrix
65| placeholder ← matrix(NA, nrow = 20, ncol = 1)
66|
67| # Run a loop to foreecast all 20 countries, using naive method
68| for (.i in 1:20) {
69|   m_naive_i ← naive(highly_skilled[, .i], h = 1)
70|   placeholder[.i] ← data.frame(unlist(m_naive_i[5]))[6]][1, 1]
71| }
72| # [5] = fitted values; [6] = 2018; [1,1] = only the numeric value
73|
74| # GEO and IDN have 2018 data, plug actual numbers back
75| placeholder[c(7, 8)] ← highly_skilled[6, c(7, 8)]
76|
77| # RUS needs separate calculation
78| # ISCED 7 = 101766 (Type 1) + 170437 (Type 2) = 272203 (total masters)
79| # ISCED 8 = 15465 (Type 1) + 330 (Type 2) = 15795 (total PhDs)
80| # Total tertiary WITHOUT PhD = 933153
81| # ⇒ Total tertiary = 933153 + 15795 = 948948
82| # highly_skilled (RUS) = (272203 + 15795) / 948948 = 0.30349187
83| placeholder[16] ← 0.30349187
84|
85| # Save results to "bookshelf"
86| fki_raw[, 2] ← placeholder * 100
87|
88| rm(
89|   isced_7, isced_8, total_tertiary,
90|   highly_skilled, placeholder, m_naive_i
91| )
92|
93| # Sub-section 2.2: Mean year of schooling
94| mean_year_of_schooling ← read.csv("mean_year_of_schooling.csv",
95|   header = F, sep = "\t"
96| )
97| fki_raw[, 3] ← mean_year_of_schooling[, 2]
98|
99| rm(mean_year_of_schooling)

```

```

100|
101|
102| # Section 3: Use
103|
104| gpea <- read.csv("gpea.csv", header = T, sep = "\t")
105| gpea <- ts(gpea, start = 2011, end = 2017, frequency = 1)
106|
107| # # Visualise data in both original and ln forms. Contain trend?
108| # pdf("../Figures/use.pdf", width = 12.94, height = 9.15)
109|
110| # # Re-set canvas layout to 2x2
111| # par(mfcol = c(2, 2))
112|
113| # # Add extra space to the right of plot area
114| # par(mar = c(5.1, 4.1, 4.1, 2.1), xpd = TRUE)
115|
116| # # Plot GPEA in original form
117| # ts.plot(gpea,
118| #   type = "b", col = 1:20,
119| #   xlab = "Year", ylab = "Percent", main = "GPEA to GDP ratio"
120| # )
121|
122| # # Remove extra gap between the two graphs
123| # par(mar = c(5.1, 4.1, 0, 2.1), xpd = TRUE)
124|
125| # # Repeat GPEA, but for the ln() version
126| # ts.plot(log(gpea),
127| #   type = "b", col = 1:20,
128| #   xlab = "Year", ylab = "ln( percent )"
129| # )
130|
131| # # Plot ICA in original form
132| # par(mar = c(5.1, 4.1, 4.1, 6.1), xpd = TRUE)
133| # ts.plot(ica,
134| #   type = "b", col = 1:20,
135| #   xlab = "Year", ylab = "Percent", main = "ICA to GDP ratio"
136| # )
137| # # Add the legend
138| # legend("topright",
139| #   inset = c(-0.2, 0), colnames(ica),
140| #   col = 1:20, lty = 1, cex = 0.875
141| # )
142|
143| # # Remove extra gap between the two graphs
144| # par(mar = c(5.1, 4.1, 0, 6.1), xpd = TRUE)
145|
146| # # Repeat, but for the ln()
147| # ts.plot(log(ica),
148| #   type = "b", col = 1:20,
149| #   xlab = "Year", ylab = "ln( percent )"
150| # )
151| # # Add the legend
152| # legend("topright",
153| #   inset = c(-0.2, 0), colnames(ica),
154| #   col = 1:20, lty = 1, cex = 1.07
155| # )
156| # dev.off()
157|
158| # Decision: since the ln() version is not flat, original time series
159| # contain trend. Use Holt method rather than simple exponential smoothing.

```

```

160 |
161 | # Run a time series forecast using Holt method
162 |
163 | # Create a placeholder matrix
164 | placeholder <- matrix(NA, nrow = 20, ncol = 1)
165 |
166 | # Run a loop to forecast all 13 countries, using Holt method
167 | for (.i in 1:20) {
168 |   m_holt_i <- holt(gpea[, .i], h = 1)
169 |   placeholder[.i] <- m_holt_i[2]
170 | } # Ignore warnings
171 |
172 | # Only keep the 2018 forecasts
173 | placeholder <- unlist(placeholder)
174 |
175 | # Run PER (#13) separately because it misses both 2017 and 2018 data
176 | m_holt_PER <- holt(gpea[, 13], h = 2); summary(m_holt_PER)
177 | placeholder[13] <- 16.02698
178 |
179 | # Push placeholder to fki_raw
180 | fki_raw[, 4] <- placeholder
181 |
182 | rm(gpea, placeholder, m_holt_i, m_holt_PER)
183 |
184 | # Sub-section 3.2: Insurance company assets (ica)
185 |
186 | ica <- read.csv("ica.csv", header = T, sep = "\t")
187 | ica <- ts(ica, start = 2011, end = 2017, frequency = 1)
188 |
189 | placeholder <- matrix(NA, nrow = 20, ncol = 1)
190 |
191 | for (.i in 1:20) {
192 |   m_holt_i <- holt(ica[, .i], h = 1)
193 |   placeholder[.i] <- m_holt_i[2]
194 | } # Ignore warnings
195 |
196 | placeholder <- unlist(placeholder)
197 |
198 | m_holt_CAN <- holt(ica[, 3], h = 2); summary(m_holt_CAN)
199 | m_holt_IND <- holt(ica[, 8], h = 2); summary(m_holt_IND)
200 | m_holt_ITA <- holt(ica[, 9], h = 2); summary(m_holt_ITA)
201 | m_holt_POL <- holt(ica[, 14], h = 2); summary(m_holt_POL)
202 | m_holt_USA <- holt(ica[, 20], h = 2); summary(m_holt_USA)
203 |
204 | placeholder[c(3, 8, 9, 14, 20)] <- c(
205 |   77.72768, 4.611597, 51.2596, 9.534750, 30.18295
206 | )
207 |
208 | fki_raw[, 5] <- placeholder
209 |
210 | rm(ica, placeholder, list = ls(pattern = "^m.holt"))
211 |
212 | # Sub-section 3.3: Individuals using the Internet (ius)
213 |
214 | ius <- read.csv("ius.csv", header = T, sep = "\t")
215 | ius <- ts(ius, start = 2009, end = 2018, frequency = 1)
216 |
217 | m_holt_CAN <- holt(ius[1:9, 3], h = 1); summary(m_holt_CAN)
218 | m_holt_CHL <- holt(ius[1:9, 4], h = 1); summary(m_holt_CHL)
219 | m_holt_USA <- holt(ius[1:9, 20], h = 1); summary(m_holt_USA)

```



```

220 |
221 | ius_2018 ← ius[10, ] # Only want 2018 data
222 | ius_2018[3] ← 93.58751 # CAN
223 | ius_2018[4] ← 89.5309 # CHL
224 | ius_2018[20] ← 84.88108 # USA
225 |
226 | fki_raw[, 6] ← ius_2018
227 |
228 | rm(list = ls(pattern = "^ius"))
229 | rm(list = ls(pattern = "^m_holt_"))
230 |
231 |
232 | # Section 4: Need
233 |
234 | # Subsection 4.1: Pension fund assets (pfa)
235 | pfa ← read.csv("pfa.csv", header = T, sep = "\t")
236 | # Delete GEO (#7) due to all missing. Will come back to it later.
237 | pfa ← ts(pfa[, -7], start = 2008, end = 2017, frequency = 1)
238 |
239 | placeholder ← matrix(NA, nrow = 19, ncol = 1)
240 |
241 | for (.i in 1:19) {
242 |   m_holt_i ← holt(pfa[, .i], h = 1)
243 |   placeholder[.i] ← m_holt_i[2]
244 | }
245 |
246 | placeholder ← unlist(placeholder)
247 |
248 | # Calculate GEO
249 | # From Georgia Pension Agency:
250 | #   2019 mesub_ind_eting minute: 372,113,933 GEL
251 | # From GeoStat website:
252 | #   2018 gdp = 44.6 billion GEL
253 |
254 | fki_raw[, 7] ← c(
255 |   placeholder[1:6],
256 |   372113934 / 44600000000 * 100, # Insert GEO figure
257 |   placeholder[7:19]
258 | )
259 |
260 | rm(pfa, placeholder, m_holt_i)
261 |
262 | # Subsection 4.2: Aggregate consumption (ac)
263 |
264 | ac ← read.csv("ac.csv", header = F, row.names = 1, sep = "\t")
265 | gdp ← read.csv("gdp.csv", header = F, row.names = 1, sep = "\t")
266 |
267 | fki_raw[, 8] ← unlist(ac * 0.02 / gdp * 100)
268 |
269 | fki_raw[, 9] ← unlist(gdp)
270 |
271 | rm(ac, gdp)
272 |
273 | # Subsection 4.3: Ageing
274 |
275 | ageing ← read.csv("ageing.csv", header = T, sep = "\t")
276 | attach(ageing)
277 | names(ageing)
278 |
279 | # Calculate total population
280 | poptotal_f ← pop0to14_f + pop15to64_f + pop65plus_f

```

```

281 | poptotal_m ← pop0to14_m + pop15to64_m + pop65plus_m
282 | # Calculate population between 15 and 19
283 | # Need to divide by 100 to get decimals
284 | pop15to19_f ← poptotal_f * per15to19_f / 100
285 | pop15to19_m ← poptotal_m * per15to19_m / 100
286 | # Calculate population between 0 and 19
287 | pop0to19_f ← pop0to14_f + pop15to19_f
288 | pop0to19_m ← pop0to14_m + pop15to19_m
289 | # Calculate population between 20 and 64
290 | pop20to64_f ← poptotal_f - pop0to19_f - pop65plus_f
291 | pop20to64_m ← poptotal_m - pop0to19_m - pop65plus_m
292 | # Calculate 64+ / 20-to-64 ratio 'GEO' = 0.419;
293 | ageing_ratio ← I(
294 |   (pop65plus_f + pop65plus_m) / (pop20to64_f + pop20to64_m)
295 | )
296 | # Split data into 2018 [ , 1] and 2009 [ , 2] portions
297 | ageing ← cbind(ageing_ratio[1:20], ageing_ratio[21:40])
298 | fki_raw[, 10] ← (ageing[, 1] - ageing[, 2]) / ageing[, 2]
299 |
300 | rm(ageing, ageing_ratio, list = ls(pattern = "^pop"))
301 |
302 |
303 | # Section 5: FKI
304 |
305 | fki_raw ← fki_raw[, -9] # Throw away gdp (already in ac)
306 | round(fki_raw, digits = 3) # Inspect data
307 |
308 | # Save data to an external file
309 | library(data.table); setDTthreads(0)
310 | fwrite(round(fki_raw, digits = 3), file = "fki_raw.csv", row.names = T)
311 |
312 | # Subection 5.0: Standardise each variable to [0.01,0.99] range
313 | fki_stand ← matrix(NA, nrow = dim(fki_raw)[1], ncol = dim(fki_raw)[2])
314 | dimnames(fki_stand) ← dimnames(fki_raw)
315 |
316 | library(scales) # I wish this function could have "by.col = T". Oh well.
317 | for (.j in 1:dim(fki_raw)[2]) {
318 |   fki_stand[, .j] ← rescale(fki_raw[, .j], to = c(0.01, 0.99))
319 | }
320 |
321 | rm(fki_raw)
322 |
323 | fki_stand ← data.frame(fki_stand)
324 | attach(fki_stand)
325 |
326 | # Subsection 5.1: Economic capacity (sub_ind_ec)
327 |
328 | sub_ind_ec ← gdp_per_capita
329 |
330 | # Subsection 5.2: Education and training (sub_ind_et)
331 |
332 | wt_highly_skilled ← 1 / sd(highly_skilled)
333 | wt_mean_year_of_schooling ← 1 / sd(mean_year_of_schooling)
334 |
335 | sub_ind_et ← (highly_skilled^wt_highly_skilled *
336 |   mean_year_of_schooling^wt_mean_year_of_schooling)^
337 |   (1 / (wt_highly_skilled + wt_mean_year_of_schooling))
338 |
339 | rm(list = ls(pattern = "^wt"))
340 |
341 | # Subsection 5.3: Use (sub_ind_use)

```

```

342|
343| sub_ind_u ← (gpea + ica)^ius
344|
345| # Subsection 5.4: Need (sub_ind_need)
346|
347| sub_ind_n ← (pfa + ac)^ageing
348|
349| ## Subsection 5.5: FKI
350|
351| wt_ec ← 1 / sd(sub_ind_ec)
352| wt_et ← 1 / sd(sub_ind_et)
353| wt_u ← 1 / sd(sub_ind_u)
354| wt_n ← 1 / sd(sub_ind_n)
355|
356| fki ← (
357|   sub_ind_ec^wt_ec *
358|   sub_ind_et^wt_et *
359|   sub_ind_u^wt_u *
360|   sub_ind_n^wt_n
361| ) ^ (
362|   1 / (wt_ec + wt_et + wt_u + wt_n)
363| )
364|
365| rm(list = ls(pattern = "^wt"))
366|
367| l3 ← data.frame(
368|   round(
369|     cbind(fki, sub_ind_ec, sub_ind_et, sub_ind_u, sub_ind_n),
370|     digits = 3
371|   )
372| )
373| rownames(l3) ← rownames(fki_stand)
374| attach(l3)
375|
376| rm(fki_stand, fki, sub_ind_ec, sub_ind_et, sub_ind_u, sub_ind_n)
377|
378| # Display country-level FKI, default by country code
379| l3
380|
381| # Sort FKI by country (highest to lowest)
382| l3_ordered ← l3[order(-fki), ]
383| l3_ordered
384| fwrite(l3_ordered, file = "fki.csv", row.names = T)
385|
386| pdf("../..//Figures/FKI.pdf")
387|   barplot(l3_ordered$fki,
388|     names.arg = rownames(l3_ordered),
389|     xlab = "Country", las = 2, ylab = "Financial Knowledge Index (FKI)",
390|     ylim = c(0, 1), main = "FKI of 20 participating countries"
391|   )
392| dev.off()

```

## B.2.4 Data Reimport

```

1| # Housekeeping
2| library(Orcs)
3| setwdOS(lin = "~/", win = Sys.getenv("USERPROFILE"))
4|
5| # Import SPSS file into R
6| library(intsvy)

```

```

7 finlit ← pisa.select.merge(
8   student.file = "CY07_MSU_FLT_QQQ.SAV", # file ext in capital
9   school.file = "CY07_MSU_SCH_QQQ.sav", # file ext in lower case
10  student = c(
11    # Control variables
12    "ST004D01T", # Student (Standardized) Gender
13    "IMMIG", # Index Immigration status
14    "ESCS", # Index of economic, social and cultural status
15    # Mediators
16    "FCFMLRTY", # Familiarity with concepts of finance (Sum)
17    "FLCONFIN", # Confidence about financial matters (WLE)
18    # Academic
19    "FLSCHOOL", # Financial education in school lessons (WLE)
20    # Safety
21    "BEINGBULLIED", # Student's experience of being bullied (WLE)
22    # Community
23    "FLFAMILY" # Parental involvement in matters of Financial Literacy (WLE)
24  ),
25  school = c(
26    "STRATIO", # Student-teacher ratio
27    "EDUSHORT", # Shortage of educational material (WLE)
28  ),
29  countries = c(
30    "BRA", "BGR", "CAN", "CHL", "EST",
31    "FIN", "GEO", "IDN", "ITA", "LVA",
32    "LTU", "NLD", "PER", "POL", "PRT",
33    "RUS", "SRB", "SVK", "ESP", "USA"
34  )
35 )
36
37 names(finlit)
38 # Throw away columns that I do not need
39 finlit ← finlit[, -c(5,7:86)] # 5 = BOOKID; 7:86 = resampling weights
40
41 # Some var need recording
42 library(car)
43
44 # Re-code Russian territories to RUS
45 finlit$CNT ← recode(finlit$CNT, "
46   'QMR' = 'RUS';
47   'QRT' = 'RUS'
48 ")
49
50 # Input country-level FKI
51 FKI ← recode(finlit$CNT, "
52   'NLD' = 0.940;
53   'USA' = 0.937;
54   'CAN' = 0.784;
55   'ITA' = 0.762;
56   'FIN' = 0.724;
57   'ESP' = 0.627;
58   'LTU' = 0.613;
59   'PRT' = 0.591;
60   'BGR' = 0.583;
61   'EST' = 0.577;
62   'SVK' = 0.559;
63   'POL' = 0.555;
64   'LVA' = 0.550;
65   'CHL' = 0.544;
66   'RUS' = 0.450;
67   'GEO' = 0.424;

```

```

68 | 'SRB' = 0.423;
69 | 'PER' = 0.309;
70 | 'BRA' = 0.141;
71 | 'IDN' = 0.122
72 | ")
73 |
74 | # Recode ST004D01T from Sex to Male
75 | MALE ← finlit$ST004D01T - 1
76 |
77 | # Recode IMMIG to 1st and 2nd generation
78 | IMMI1GEN ← recode(finlit$IMMIG, "
79 |   1 = 0;
80 |   2 = 0;
81 |   3 = 1
82 | ")
83 |
84 | IMMI2GEN ← recode(finlit$IMMIG, "
85 |   1 = 0;
86 |   2 = 1;
87 |   3 = 0
88 | ")
89 |
90 | # Revert coding direction: bigger number ⇒ safer school
91 | NOBULLY ← finlit$BEINGBULLIED * (-1)
92 |
93 | # Stitch spreadsheet together
94 | names(finlit)
95 | finlit ← cbind(FKI, finlit[, c(2:35)], MALE, IMMI1GEN, IMMI2GEN, finlit[, c(38:50)])
96 |
97 | # Remove cases whose school weights (col #48) are NA
98 | obs0 ← dim(finlit)[1]
99 | finlit ← finlit[complete.cases(finlit[, 48]), ]
100 | obs1 ← dim(finlit)[1]
101 | obs0 - obs1 # 12 cases contained missing school weights and have been dropped
102 | rm(obs0, obs1)
103 |
104 | # Use data.table for better RAM management
105 | library(data.table); setDTthreads(0) # 0 means all the available cores
106 | # Export data into a CSV file for faster import next time
107 | fwrite(finlit, file = "finlit.csv", na = "NA", row.names = F, col.names = T)

```



# Appendix C Derivation of Moderated Mediation Effect

## C.1 Models with Mediators Only

Consider a SEM model shown in [Figure C.1](#) (excluding any paths in green), where

$$\begin{cases} Y = \mu_0 + b_1 M_1 + b_2 M_2 + c_1 X_1 + c_2 X_2 + c_3 X_3 \\ M_1 = \mu_1 + a_{11} X_1 + a_{21} X_2 + a_{31} X_3 \\ M_2 = \mu_2 + a_{12} X_1 + a_{22} X_2 + a_{32} X_3 \end{cases}$$

or, in matrix form

$$\begin{cases} Y = \mu_0 + \mathbf{b}^\top \mathbf{m} + \mathbf{c}^\top \mathbf{x} \\ \mathbf{m} = \boldsymbol{\mu} + \mathbf{A}^\top \mathbf{x} \end{cases} \quad (\text{C.1})$$

where

$$\mathbf{x}_{3 \times 1} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \quad \mathbf{m}_{2 \times 1} = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \quad \mathbf{b}_{2 \times 1} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \mathbf{c}_{3 \times 1} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}, \quad \boldsymbol{\mu}_{2 \times 1} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ and } \mathbf{A}_{3 \times 2} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

[Equation \(C.1\)](#) can be written as a total equation:

$$Y = \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} + \mathbf{b}^\top \mathbf{A}^\top \mathbf{x} + \mathbf{c}^\top \mathbf{x} = \left( \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} \right) + \mathbf{x}^\top (\mathbf{A} \mathbf{b} + \mathbf{c}) \quad (\text{C.2})$$

where  $\mu_0 + \mathbf{b}^\top \boldsymbol{\mu}$  is the intercept,  $\mathbf{A} \mathbf{b}$  is the indirect effect and  $\mathbf{c}$  is the direct effect.

## C.2 Models with Moderated Mediators

Now introduce two moderators  $D_1$  and  $D_2$  (green paths in [Figure C.1](#)).

In scalar notation:

$$\begin{aligned}
Y_{\text{mod}} = & \mu_0 + b_1 M_1 + b_2 M_2 + c_1 X_1 + c_2 X_2 + c_3 X_3 \\
& + f_1 D_1 + f_2 D_2 \\
& + g_{11} X_1 D_1 + g_{12} X_1 D_2 \\
& + g_{21} X_2 D_1 + g_{22} X_2 D_2 \\
& + g_{31} X_3 D_1 + g_{32} X_3 D_2 \\
& + h_{11} M_1 D_1 + h_{12} M_1 D_2 \\
& + h_{21} M_2 D_1 + h_{22} M_2 D_2
\end{aligned}$$

and in matrix notation:

$$Y_{\text{mod}} = \mu_0 + \mathbf{b}^\top \mathbf{m} + \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{d} + \text{tr}(\mathbf{G}^\top \mathbf{x} \mathbf{d}^\top) + \text{tr}(\mathbf{H}^\top \mathbf{m} \mathbf{d}^\top) \quad (\text{C.3})$$

where,

$$\mathbf{f}_{2 \times 1} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \quad \mathbf{d}_{2 \times 1} = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}, \quad \mathbf{G}_{3 \times 2} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \end{pmatrix}, \quad \mathbf{H}_{2 \times 2} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix},$$

and  $\text{tr}(\cdot)$  is the trace operator.

Since  $\mathbf{m} = \boldsymbol{\mu} + \mathbf{A}^\top \mathbf{x}$ , Equation (C.3) can be expanded into:

$$\begin{aligned}
Y_{\text{mod}} = & \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} + \mathbf{b}^\top \mathbf{A}^\top \mathbf{x} + \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{d} + \text{tr}(\mathbf{G}^\top \mathbf{x} \mathbf{d}^\top) + \text{tr}(\mathbf{H}^\top \boldsymbol{\mu} \mathbf{d}^\top) + \text{tr}(\mathbf{H}^\top \mathbf{A}^\top \mathbf{x} \mathbf{d}^\top) \\
= & \left[ \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} + \mathbf{f}^\top \mathbf{d} + \text{tr}(\mathbf{H}^\top \boldsymbol{\mu} \mathbf{d}^\top) \right] + \left[ (\mathbf{b}^\top \mathbf{A}^\top + \mathbf{c}^\top) \mathbf{x} + \text{tr}(\mathbf{d}^\top (\mathbf{G}^\top + \mathbf{H}^\top \mathbf{A}^\top) \mathbf{x}) \right] \\
= & \left[ \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} + \mathbf{f}^\top \mathbf{d} + \text{tr}(\mathbf{H}^\top \boldsymbol{\mu} \mathbf{d}^\top) \right] + \left[ (\mathbf{b}^\top \mathbf{A}^\top + \mathbf{c}^\top) \mathbf{x} + \mathbf{d}^\top (\mathbf{G}^\top + \mathbf{H}^\top \mathbf{A}^\top) \mathbf{x} \right] \\
= & \left[ \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} + \mathbf{f}^\top \mathbf{d} + \text{tr}(\mathbf{H}^\top \boldsymbol{\mu} \mathbf{d}^\top) \right] + \mathbf{x}^\top [\mathbf{A} \mathbf{b} + \mathbf{c} + \mathbf{G} \mathbf{d} + \mathbf{A} \mathbf{H} \mathbf{d}] \\
= & \left[ \mu_0 + \mathbf{b}^\top \boldsymbol{\mu} + \mathbf{f}^\top \mathbf{d} + \text{tr}(\mathbf{H}^\top \boldsymbol{\mu} \mathbf{d}^\top) \right] + \mathbf{x}^\top [\mathbf{A} (\mathbf{b} + \mathbf{H} \mathbf{d}) + (\mathbf{c} + \mathbf{G} \mathbf{d})] \quad (\text{C.4})
\end{aligned}$$

Equation (C.4) differs from Equation (C.2) by one extra term  $\mathbf{f} \mathbf{d}^\top + \text{tr}(\mathbf{H}^\top \boldsymbol{\mu} \mathbf{d}^\top)$  in the intercept. The indirect effect  $\mathbf{A} \mathbf{b}$  expanded to  $\mathbf{A} (\mathbf{b} + \mathbf{H} \mathbf{d})$  as a result of introducing the moderators and the direct effect grows from  $\mathbf{c}$  to  $\mathbf{c} + \mathbf{G} \mathbf{d}$ .



Expand the indirect and direct effects back to their scalar forms:

indirect effects

$$\begin{aligned}
&= \mathbf{A} (\mathbf{b} + \mathbf{H}\mathbf{d}) \\
&= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \left[ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \right] \\
&= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} b_1 + h_{11}D_1 + h_{12}D_2 \\ b_2 + h_{21}D_1 + h_{22}D_2 \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_1 + a_{11}h_{11}D_1 + a_{11}h_{12}D_2 + a_{12}b_2 + a_{12}h_{21}D_1 + a_{12}h_{22}D_2 \\ a_{21}b_1 + a_{21}h_{11}D_1 + a_{21}h_{12}D_2 + a_{22}b_2 + a_{22}h_{21}D_1 + a_{22}h_{22}D_2 \\ a_{31}b_1 + a_{31}h_{11}D_1 + a_{31}h_{12}D_2 + a_{32}b_2 + a_{32}h_{21}D_1 + a_{32}h_{22}D_2 \end{pmatrix};
\end{aligned}$$

direct effects

$$\begin{aligned}
&= \mathbf{c} + \mathbf{G}\mathbf{d} \\
&= \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \end{pmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \\
&= \begin{pmatrix} c_1 + g_{11}D_1 + g_{12}D_2 \\ c_2 + g_{21}D_1 + g_{22}D_2 \\ c_3 + g_{31}D_1 + g_{32}D_2 \end{pmatrix}.
\end{aligned}$$

### C.3 Mplus Execution

The **DEFINE:** and **MODEL:** sections of the Mplus code is given as following:

```

1 DEFINE:
2
3     ! G matrix
4     X1D1 = X1 * D1;
5     X2D1 = X2 * D1;
6     X3D1 = X3 * D1;
7     X1D2 = X1 * D2;
8     X2D2 = X2 * D2;
9     X3D2 = X3 * D2;

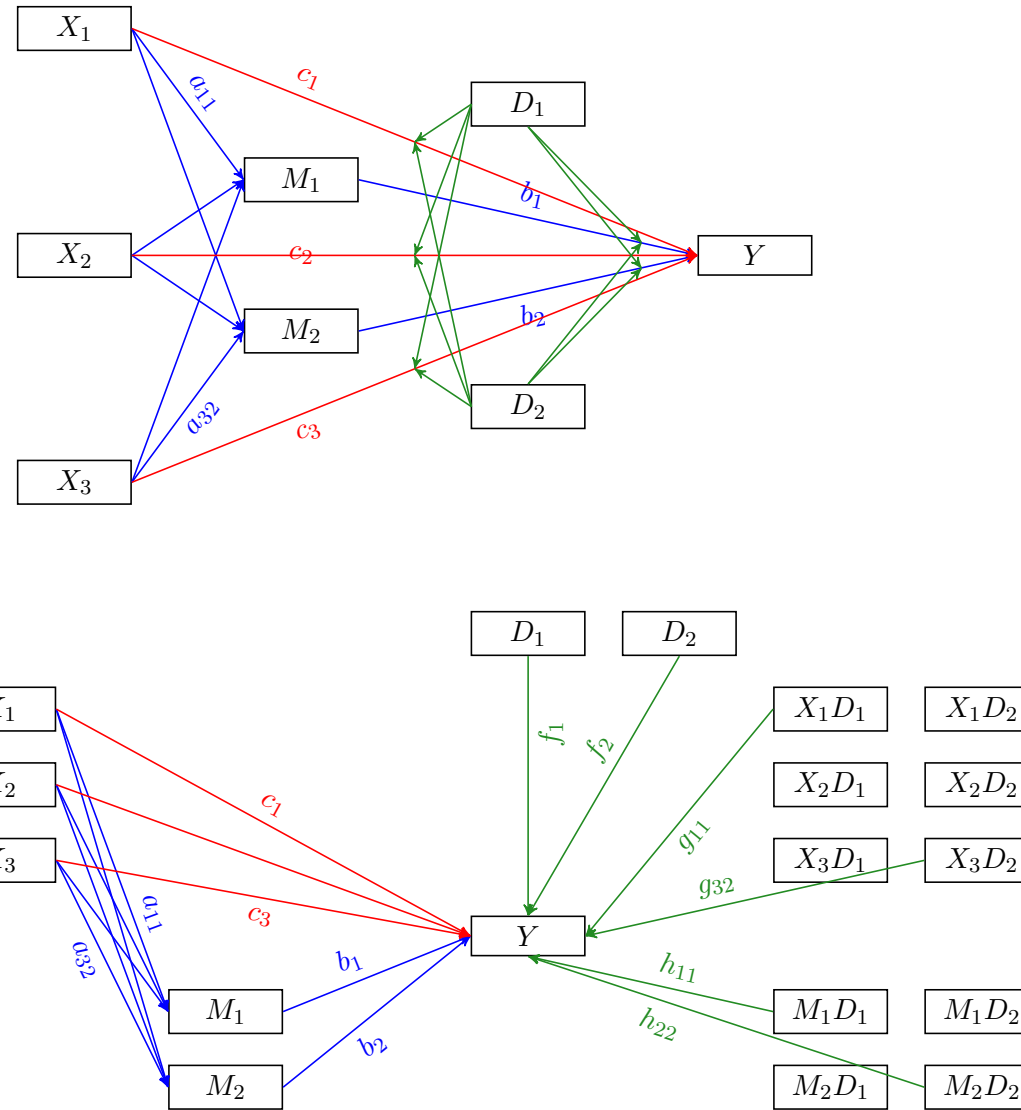
```

```

10      ! H matrix
11      M1D1 = M1 * D1;
12      M2D1 = M2 * D1;
13      M1D2 = M1 * D2;
14      M2D2 = M2 * D2;
15
16 MODEL:
17
18      [Y] (mu0);
19      Y on M1 (b1);
20      Y on M2 (b2);
21      ! ---
22      Y on M1D1 (h11);
23      Y on M2D1 (h21);
24      Y on M1D1 (h12);
25      Y on M2D1 (h22);
26      ! ---
27      Y on X1 (c1);
28      Y on X2 (c2);
29      Y on X3 (c3);
30      ! ---
31      Y on D1 (f1);
32      Y on D2 (f2);
33      ! ---
34      Y on X1D1 (g11);
35      Y on X2D1 (g21);
36      Y on X3D1 (g31);
37      Y on X1D2 (g12);
38      Y on X2D2 (g22);
39      Y on X3D2 (g32);
40
41      [M1] (mu1);
42      M1 on X1 (a11);
43      M1 on X2 (a21);
44      M1 on X3 (a31);
45
46      [M2] (mu2);
47      M2 on X1 (a12);
48      M2 on X2 (a22);
49      M2 on X3 (a32);
50

```

**Figure C.1**  
*Moderated Mediation Model*



*Note.* A moderated mediation is shown in both model diagram (upper panel) and statistical diagram (lower panel). Direct paths, indirect paths and moderations are differentiated by colour.

