

**Latin American Laboratory for Assessment of the Quality
of Education - LLECE**

**Using Stata to assess the
achievement of Latin American
students in Mathematics,
Reading and Science**

Roy Costilla

Outline

- 1. Why Stata?**
- 2. What the SERCE is?**
- 3. Stata at work**
- 4. Challenges**
- 5. Concluding remarks**

1. Why Stata?

- **Managing Complex Designs**
 - Weights, strata, psu's, fpc, etc.
- **Alternative variance estimation methods: Taylor linearization, Replication Methods and Bootstrap**
- **Matrix Language (Watson, 2005)**
 - Allows you to store estimation results
- **Programming and Macros**
 - Allows you to automate the whole estimation and testing process.

2. What the SERCE is?

- **Second Regional Comparative and Explanatory Study (OREALC/UNESCO Santiago, 2008)**
- **Objective: Give insight into the learning acquired by Latin American and Caribbean students and analyze the associated factors related to that learning.**
 - **Primary school students who during the period 2005 /2006 attended third and sixth grades**
 - **Areas of Mathematics, Language (Reading and Writing) and Natural Science.**
- **Collective effort of the National Assessment Systems in Latin America and the Caribbean, articulated by the Laboratory for Assessment of the Quality of Education (LLECE).**

Participants



2. What the SERCE is?. Instruments

◆ Tests:

- **Asses conceptual domains and cognitive processes.**
- **Based on common curricular elements (OREALC/UNESCO Santiago, 2005) and the life-skills approach (Delors et al. ,1996)**
- **IRT to asses students' ability**
- **Items:**
 - **4 Levels of Performance**
 - **Balanced incomplete blocks of Items.**
 - **Close and open-ended questions**

◆ Questionnaires

- **Students, teachers, principals, and parents.**

2. What the SERCE is?. Design

- **Stratification:**
 - **3 Domains: Rural, Urban Public, Urban Private**
 - **Aprox. 14 Strata on each country**
- **Clustered Sampling:**
 - **Simple random sample (SRS) of schools (PSU's) without replacement**
 - **All third and sixth grade students on each selected school**

Schools	Classrooms		Students	
	3rd	6th	3rd	6th
3.065	4.627	4.227	100.752	95.288

- **The design is approximated by a two-stage stratified design with PSUs sampled with replacement**

2. What the SERCE is?. Design and...

- **Weights:**
 - Take into account unequal probabilities of selection, stratification, clustering, non-response and undercoverage
- **Taylor linearization to estimate variance (Wolter, 1985; Shao, 1996; Judkins, 1990; Kreuter & Valliant, 2007)**
 - + No Computationally intensive
 - Releasing of the unit identifiers in public data sets
- **SERCE's first report:**
 - Mean scores and Proportions and Hypothesis Testing.
- **Databases and technical documentation will be publicly available in 2009/1**

3. Stata at work. Database

```
. use m3, clear
```

```
. describe
```

Contains data from **m3.dta**

obs: 96.663

vars: 92

size: 15.466.080 (98,4% of memory free)

7 Nov 2008 20:20

variable name	storage type	display format	value label	variable label
pais	str2	%2s		País
centro_educat~o	str3	%3s		Centro educativo
grado	str1	%1s		Grado
aula	str2	%2s		Aula
area	str1	%1s		Área
estudiante	str2	%2s		Estudiante
cuadernillo	str2	%2s		Cuadernillo
bloque_primera	str1	%1s		Bloque primera
bloque_segunda	str1	%1s		Bloque segunda
LlavePaisCentro	str5	%5s		Identificador del país
id_grado	str6	%6s		Identificador de grado
id_gradoaula	str8	%8s		Identificador de aula
id_alumno	str10	%10s		Identificador de alumno
admrur	byte	%12.0f	admrur	Primera variable de est
estrato	long	%12.0f	estrato	Estratificacion aprigri

3. Stata at work. Declaring Complex Design

```
. svydescribe
```

```
Survey: Describing stage 1 sampling units
```

```
    pweight: peso_estudiante  
          VCE: linearized  
Single unit: missing  
Strata 1: estratoregional  
    SU 1: LlavePaisCentro  
    FPC 1: <zero>
```

Stratum	#Units	#Obs	#Obs per Unit		
			min	mean	max
10000000	9	167	9	18,6	25
10000000	31	1591	29	51,3	100
10000000	13	1230	70	94,6	133
10000000	8	210	15	26,3	43
10000000	12	748	30	62,3	95
10000001	24	286	3	11,9	30
10000001	5	234	20	46,8	116
10000001	3	61	8	20,3	36
10000001	7	330	23	47,1	72
10000001	4	512	63	128,0	233
10000001	4	84	6	21,0	35
10000001	4	213	38	53,3	71

3. Stata at work. Means

```
. svy, subpop(serce) : mean puntaje_escal_a_m3
(running mean on estimation sample)
```

Survey: Mean estimation

Number of strata =	241	Number of obs =	91252
Number of PSUs =	2686	Population size =	10907247
		Subpop. no. obs =	91252
		Subpop. size =	10907247
		Design df =	2445

	Mean	Linearized Std. Err.	[95% Conf. Interval]	
puntaje_es~3	505,1089	2,404318	500,3942	509,8236

Note: 13 strata omitted because they contain no subpopulation members.

3. Stata at work. Proportions

```
. svy, subpop(serce): proportion (nivel)
(running proportion on estimation sample)
```

Survey: Proportion estimation

```
Number of strata =      241      Number of obs      =      91252
Number of PSUs   =      2686     Population size     =     10907247
Subpop. no. obs  =      91252
Subpop. size     =     10907247
Design df       =      2445
```

```
_prop_1: nivel = <I
```

	Proportion	Linearized Std. Err.	[95% Conf. Interval]	
nivel				
_prop_1	,1122651	,0057282	,1010324	,1234978
I	,1430188	,0044257	,1343402	,1516973
II	,282574	,0044472	,2738533	,2912946
III	,3602692	,0078467	,3448823	,375656
IV	,101873	,0043448	,0933531	,1103929

Note: 13 strata omitted because they contain no subpopulation members.

3. Stata at work

➤ Perform hypothesis testing and store results

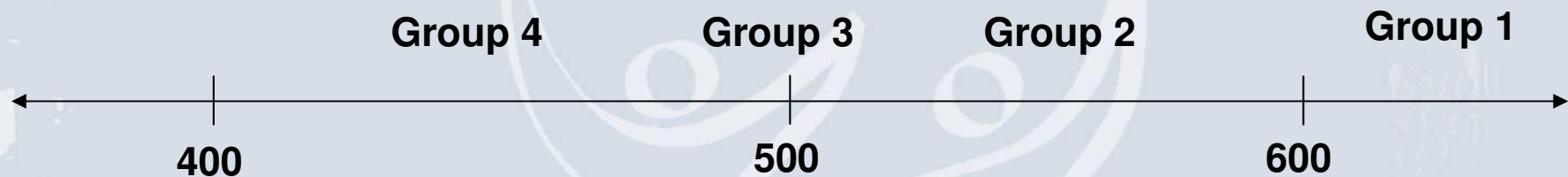
```
. svy, subpop(serce) : mean puntaje_escal_a_m3, over(rural)
. lincom [puntaje_escal_a_m3]Rural - [puntaje_escal_a_m3]Urbana
( 1) - [puntaje_escal_a_m3]Urbana + [puntaje_escal_a_m3]Rural = 0
```

	Coef.	Std. Err.	t	P> t	[95% Conf.
(1)	-51,08027	3,817521	-13,38	0,000	-58,56618

```
. matrix define Rural = ( r(estimate), r(se) , r(estimate)/r(se) , 2
. mat colnames Rural = Coef se t P_value df
. mat rownames Rural = Total
. mat list Rural
Rural[1,5]
Total   Coef      se      t      P_value      df
Total  -51,080273  3,8175211  -13,380482  1,855e-39  2445
```

3. Stata at work

- **Automation of the estimation and testing process**
 - To classify countries into groups according to its difference with the region's mean (500 with 100 of standard deviation)



- **Bonferroni's Test**
 - For each country: Test country mean score against other countries means
 - In Reading 6th aprox. $17 \times 17 = 289$ test to be performed

Mean scores comparison Reading, 6th grade

	Promedio países	Argentina	Brasil	Chile	Colombia	Costa Rica	Cuba	Ecuador	El Salvador	Guatemala	México	Nicaragua	Panamá	Paraguay	Perú	R. Dominicana	Uruguay	Nuevo León
Argentina	=		=	▼	=	▼	▼	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▼
Brasil	▲	=		▼	=	▼	▼	▲	▲	▲	=	▲	▲	▲	▲	▲	▼	▼
Chile	▲	▲	▲		▲	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	=	=
Colombia	▲	=	=	▼		▼	▼	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▼
Costa Rica	▲	▲	▲	▲	▲		▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Cuba	▲	▲	▲	▲	▲	▲		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Ecuador	▼	▼	▼	▼	▼	▼	▼		▼	=	▼	▼	▼	=	▼	▲	▼	▼
El Salvador	▼	▼	▼	▼	▼	▼	▼	▲		▲	▼	▲	▲	▲	=	▲	▼	▼
Guatemala	▼	▼	▼	▼	▼	▼	▼	=	▼		▼	▼	▼	=	▼	▲	▼	▼
México	▲	▲	=	▼	▲	▼	▼	▲	▲	▲		▲	▲	▲	▲	▲	▼	▼
Nicaragua	▼	▼	▼	▼	▼	▼	▼	▲	▼	▲	▼		=	▲	=	▲	▼	▼
Panamá	▼	▼	▼	▼	▼	▼	▼	▲	▼	▲	▼	=		▲	=	▲	▼	▼
Paraguay	▼	▼	▼	▼	▼	▼	▼	=	▼	=	▼	▼	▼		▼	▲	▼	▼
Perú	▼	▼	▼	▼	▼	▼	▼	▲	=	▲	▼	=	=	▲		▲	▼	▼
R. Dominicana	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼		▼	▼
Uruguay	▲	▲	▲	=	▲	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲		=
Nuevo León	▲	▲	▲	=	▲	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	=	

4. Challenges

- **Alternative Variance estimation methods**

- **Multilevel analysis**
 - **There is a first regional analysis**
 - **Country specific analysis**

- **LLECE and SERCE:**
 - **SERCE “pilot” of the Third study**
 - **Human resources, facilities and funding restrictions**
 - **LLECE network of the National Evaluation Systems**

5. Concluding remarks

- **We have presented the estimation of the main results of the first report of the SERCE**
 - **SERCE:**
Assessment of the performance in the domains of Mathematics, Reading and Science of third and sixth grades students in sixteen countries of Latin America and the Caribbean in 2005/2006.
 - **Mean scores and their variability by country, areas, grades and some subpopulations.**
 - **Comparisons made in order to check for the differences in performance.**

5. Concluding remarks

- **Stata's good properties to analyze survey data.**
 - **Take in to account important aspects of a complex survey design**
 - **Availability of alternative variance estimation methods.**
 - **Automation the whole estimation and testing process using matrix and macro language Stata**

References

- Delors, J. ; et.al (1996), Learning: The Treasure Within. Report to UNESCO of the International Commission on Education for the Twenty-first Century
- Frauke Kreuter & Richard Valliant, 2007. "A survey on survey statistics: What is done and can be done in Stata," Stata Journal, StataCorp LP, vol. 7(1), 1-21
- Judkins, D. (1990). Fay's Method for Variance Estimation. Journal of Official Statistics, 6,223-240
- OREALC/UNESCO Santiago (2005), Second Regional Comparative and Explanatory Study (SERCE). Curricular analysis
- OREALC/UNESCO Santiago (2008), Student achievement in Latin America and the Caribbean. Results of the Second Regional Comparative and Explanatory Study (SERCE)
- Shao, J. (1996). Resampling Methods in Sample Surveys (with Discussion). Statistics, 27,203-254
- Watson, I. (2005), 'Further processing of estimation results: Basic programming with matrices', The Stata Journal, 5(1), 83-91
- Wolter, K.M. (1985), Introduction to Variance Estimation



Thanks for you attention!

roycostilla@gmail.com

<http://llece.unesco.cl/ing/>