Sample Design Issues in PISA MEXICO

Ismael Flores Cervantes
Gustavo Flores Vázquez

September 14, 2009
Kiel, Germany

What is PISA?

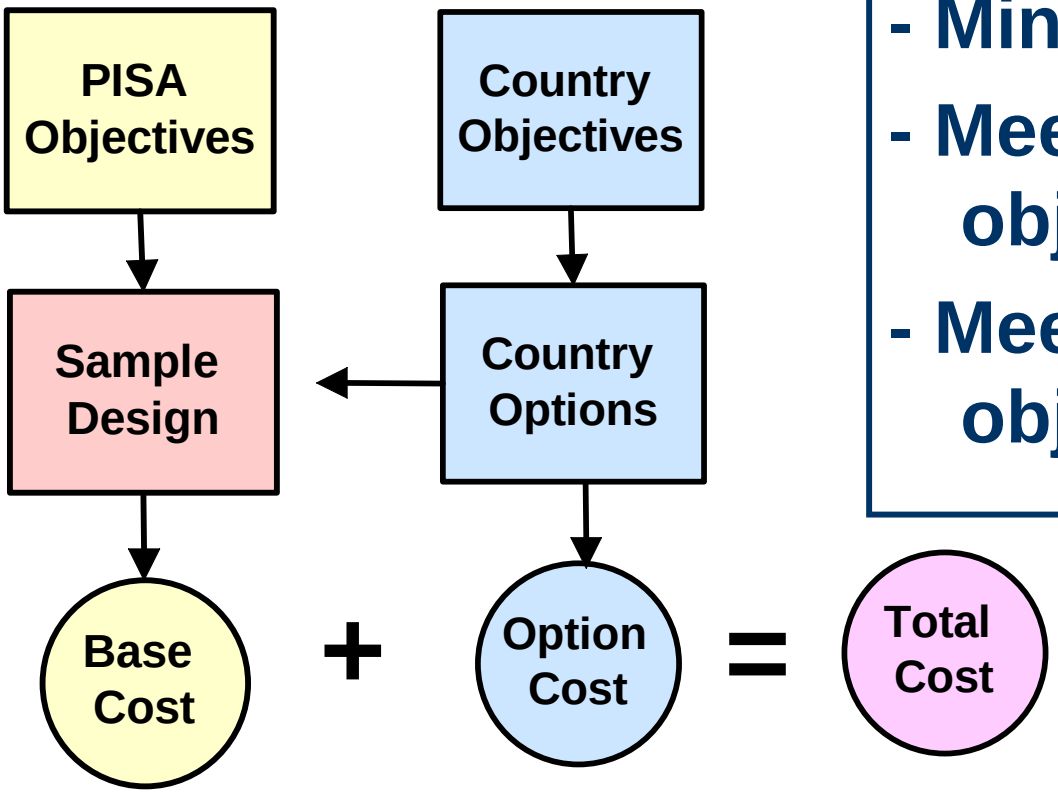
OECD's Programme for International Student Assessment (PISA) surveys

- **Survey of 15-year-old students in grade 7 and higher**
- **Every three years since 2000**
- **Complex design**
  - **Stratified two-stage with probability proportional to size (PPS) design.**
- **Use schools as clusters (PSUs)**
  - **Fixed within school sample to achieve self-weighting samples**

PISA Research Opportunities

- **Cyclical nature**
  - **Previous results can be used to improve sample design for following cycle**
- **Provide information for decision making for future cycles**
  - **Minimum sample size and associated errors**
  - **Minimum detectable differences**

Sample Design Objectives



Efficient design

- **Minimize total cost**
- **Meets PISA objective**
- **Meets country objectives**

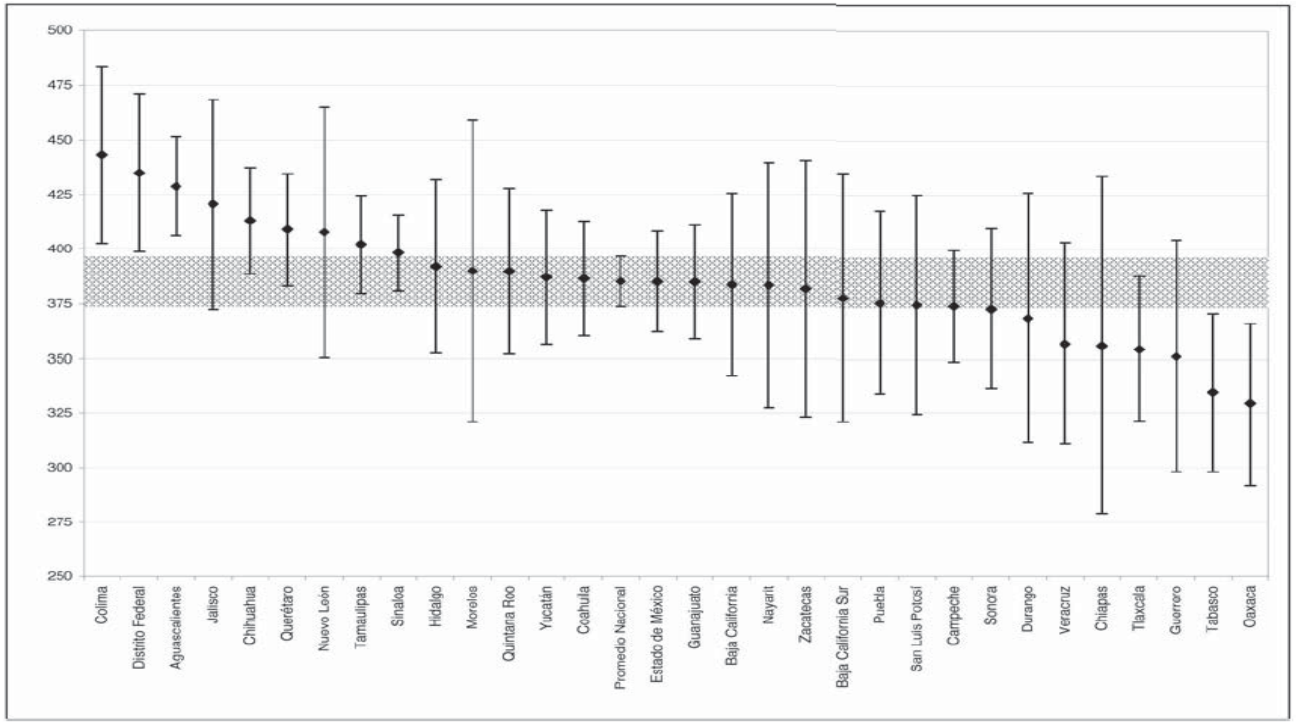
Mexico Sample

- National option (expanded sample) for estimates at state level since 2003
  - Increase the sample size by 10 times
- It did not meet national option objectives in 2003 and 2006
- Limited usability of state level data
- No corrective action in 2006
- Better in 2009, there are still improvements to made
- International objectives have been met

Implications of sample allocation

- **Limited usability of data at state level**

GRÁFICA 4.5. COMPARACIÓN DE LAS MEDIAS POR ENTIDADES EN LA ESCALA GLOBAL DE MATEMÁTICAS EN PISA 2003



Mexico's Design Characteristics

- **Frame information**
  - **Outdated (lag of one year or more)**
  - **Different sources (school's systems) and time periods**
  - **Not reliable numbers for 15 year olds in some schools**
    - **Misclassification in stratification by school size**
    - **Incorrect measure of size used in PPS sampling**

Design Effects and Effective Sample

- **Design effect**
  - Measure of efficiency of the sample design
  - Large design effect → inefficient design
  - More sample to obtain the same precision
  - Higher cost for same precision
- **Effective sample size: available for inferences**

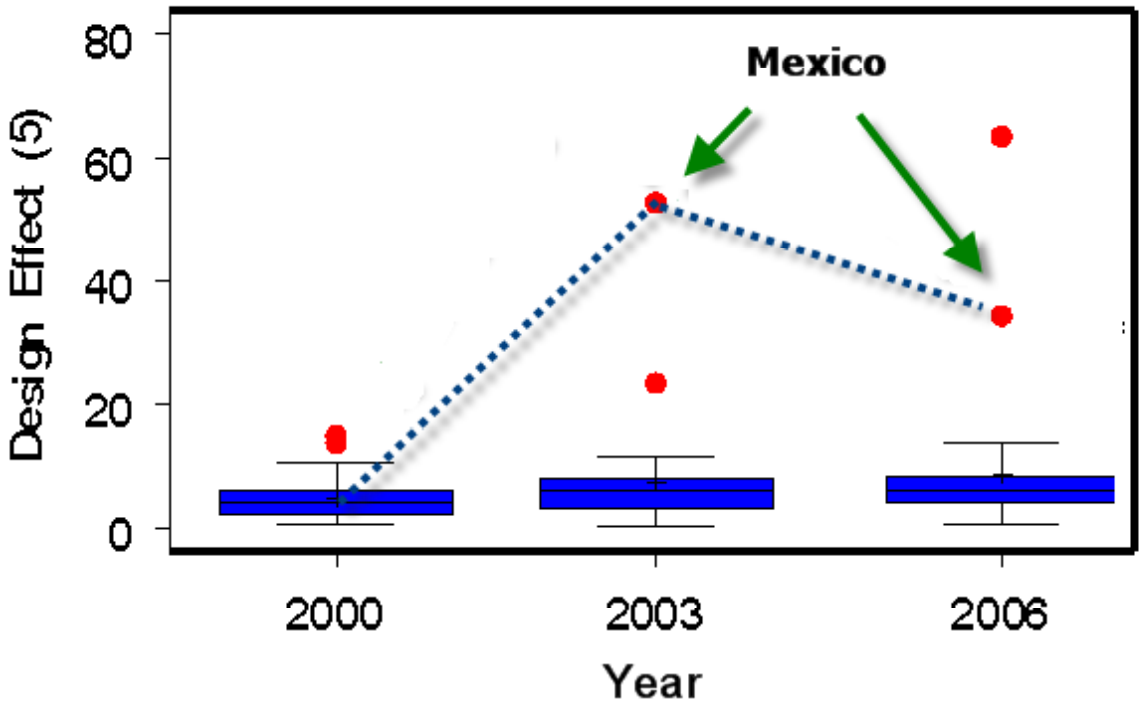
$$n_{effective} = \frac{n_{nominal}}{DEFF}$$

n nominal	Design effect	n effective
1,000	2	500
2,000	4	500
1,000	4	250

Factors that Affect the Design Effect

- Sampling unit (school)
- Stratification / sample allocation (subsampling)
- Measure of size
- Subsampling within school (students)
- Nonresponse adjustments

Total Design Effect 2000-2006



YEAR	Total Design Effect		Ratio
	Mean	Mexico	
2000	5	5	1
2003	8	53	10
2006	9	34	7

10

DEFF from Sample allocation

Y	School Size	Number of students	Sampled students	Relative Sampling	DEFF		
					A	B	Total

From PISA manual...

procedure was recommended that assumes identifying strata of both very small and moderately small schools. The underlying idea is to under-sample by a factor of two the very small school stratum and to increase proportionally the sizes of the large school strata. When there was just a single small school stratum, the procedure was modified by ignoring the parts concerning very small schools. The formulae

- Step 1: From the complete sampling frame, find the proportions of total *ENR* that come from very small schools (*P*), moderately small schools (*Q*), and larger schools (those with *ENR* of at least *TCS*) (*R*). Thus, $P + Q + R = 1$.
- Step 2: Calculate the figure *L*, where $L = 1.0 + (P/2)$. Thus *L* is a positive number slightly more than 1.0.
- Step 3: The minimum sample size for larger schools is equal to 150 × R × L, rounded to the nearest integer. It may need to be enlarged because of national considerations, such as the need to achieve minimum sample sizes for geographic regions or certain school types.
- Step 4: Calculate the mean value of *ENR* for moderately small schools (*MENR*), and for very small

Incorrect MOS and SRS

- SRS of schools in small and medium school strata (30% of population)
- SRS clusters is very inefficient with cluster (school variable size) (Cochran, 1977)

$$F_{\substack{\text{SRS}\\ \text{Cluster}}} = 1 + \frac{\sum_{c=1}^{C}\left(N_c - \bar{N}\right)N_c\bar{Y}_c}{C-1}\frac{1}{\bar{N}S_Y^2}$$

Increase in DEFF		
Year	**Small schools**	**Medium schools**
2000	**10**	**12**
2003	**18**	**28**
2006	**16**	**8**

Incorrect MOS and PPS

- PPS is used in 70% of population
- Analysis looks at effect of incorrect MOS and treats it as a form of oversampling

$$F_{\substack{\text{Error}\\ \text{in MOS}}} = \frac{1}{N^2}\sum_{c=1}^{C}\frac{N_c^2}{p_c}$$

YEAR	Coefficient of variation of weights	F MOS error
2000	40	1.2
2003	230	1.9
2006	188	1.8

A Better MOS

- PISA sampling manual mentions the use of number of students in modal grade instead of number of 15 year old students
- It has no been implemented despite differences between MOS and observed enrollment in all cycles

Modeling the School MOS

- Linear regression of transformed variables by modality and support
- We evaluate the reduction of variability of the difference as

$$R = \frac{V\left(Model\ MOS - Observed\right)}{V\left(MOS - Observed\right)}$$

- Large reduction $\rightarrow$ low ratio

Modeling the School MOS (continued)

Group	Type	R	Variables
High School	Private	59 %	# in modal grade, # total in school, # 15 years old students
	Public	43 %	# modal grade, # total in school, # 15 years old students Urban status Type (i.e., tecnologico)
Middle School	Private	43 %	# in modal grade, # total in school, Type (i.e., telesecundaria)
	Public	50 %	# modal grade, # total in school, # 15 years old students

Better Sample Allocation

	State	Size	Frame				Sample		Rel. smp rates		deff	Effective sample
			# Schools		# Students		Schools	Students				
6	B. CALIFORNI	Sn					5	36	0.5	1.0	1.03	1,261
5	B. CALIFORNI	Me					3	75	0.7	1.5		
4	B. CALIFORNI	La					34	1,190	1.0	2.0		
	Total						42	1,301				
6	ZACATECAS	Sn					11	43	0.1	1.0	1.87	642
5	ZACATECAS	Me					6	145	0.8	6.1		
4	ZACATECAS	Large	105	10%	10,768	68%	29	1,015	1.0	7.3		
	Total		1,038	100%	15,945	100%	46	1,203				

Very different sample sizes

Sample Allocation

Formula in PISA manual

- **Close to the best allocation when small schools are subsampled by a factor of 1/2**
- **Produces large effects when lower subsampling factors are implemented**

Optimization problem: allocate sample that

- **Minimize DEEF with these conditions**
  - **Subsampling factor < 1/k (reduce small schools)**
  - **Number of school less than allocated by PISA or**
  - **Number of students less than allocated by PISA**
  - **Minimum number of schools in strata > S**
  - **Effective sample size is fixed (i.e., 1,000)**

Better Sample Allocation

Sample allocation

	State	Size	Frame		Sample		Rel. smp rates		deff	Effective sample
			# Schools	# Students	Schools	Students				
06	B. CALIFORNIA	Small			4	29			1.03	1,000
05	B. CALIFORNIA	Medium			4	100				
04	B. CALIFORNIA	Large			26	900				
	Total				34	1,029				
96	ZACATECAS	Small			26	100			1.20	1,000
95	ZACATECAS	Medium			6	149				
94	ZACATECAS	Large			27	954				
	Total				59	1,203				

	Option A	2,504	1,903
	Option B	[illegible]	[illegible]
	Total	272	-97

Alternative Design Evaluation

Allocation	Number of schools	Number of students	%
2009 Alternative	1,442 1,306	40,264 31,051	77.1%
Difference	136	9,213	

Summary and Recommendations

- Cyclical nature of PISA offers opportunities to improve the sample designs for future cycles
- Need of QC procedures to track high values of design effects to avoid the 2003 and 2006 situation
- Review of sample allocation rules when lower subsampling is implemented
- Consider use of modal grade or total number of students as proxy to total number of 15 years old to reduce variance

Summary (continued)

- Explore reduction of variance by changing the sampling rate within schools
- Countries should get involved to improve design of national options. This can translate into considerable savings
- Objectives of national options should be clearly defined so the sample design can be worked ensuring that the PISA objectives are always met

Contact Information

Ismael Flores Cervantes

ismaelflorescervantes@westat.com

Gustavo Flores Vázquez

gflores@inee.edu.mx

- **Review documents: reports tables with errors**

Number of Sampled Students

School ID	State	2009		Altervative	
		Sampled Students	Student weight	Sampled Students	Student weight
12	Aguascalientes	35	14.5	35	7.2
11	Aguascalientes	35	10.7	40	9.4
8	Aguascalientes	35	8.7	40	7.6
1	Aguascalientes	35	7.9	38	7.2
14	Aguascalientes	35	7.8	38	7.2
6	Aguascalientes	35	8.0	39	7.2
21	Aguascalientes	35	6.9	33	7.2
19	Aguascalientes	35	8.9	40	7.8
...	...	...	...	...	...
7	Aguascalientes	35	7.7	38	7.2
5	Aguascalientes	35	9.3	40	8.2
20	Aguascalientes	35	3.7	20	6.5
3	Aguascalientes	35	5.5	27	7.2
18	Aguascalientes	29	5.3	26	7.2
Total		825		825	
Ratio CV	42.31%				

Optimization Problem

- Chose n students within school such that it minimizes the variation of student weights in the stratum
- Upper and lower bonds
  - Maximum 40 students
  - Minimum 25 students
  - Same total number of students as initially allocated

- **Do not use**

PISA Research Conference 2009

- A particular focus of the conference will be on issues concerning the quality and improvement of the assessment, and an important outcome will be to identify issues that may feed into an agenda for future research and development activities

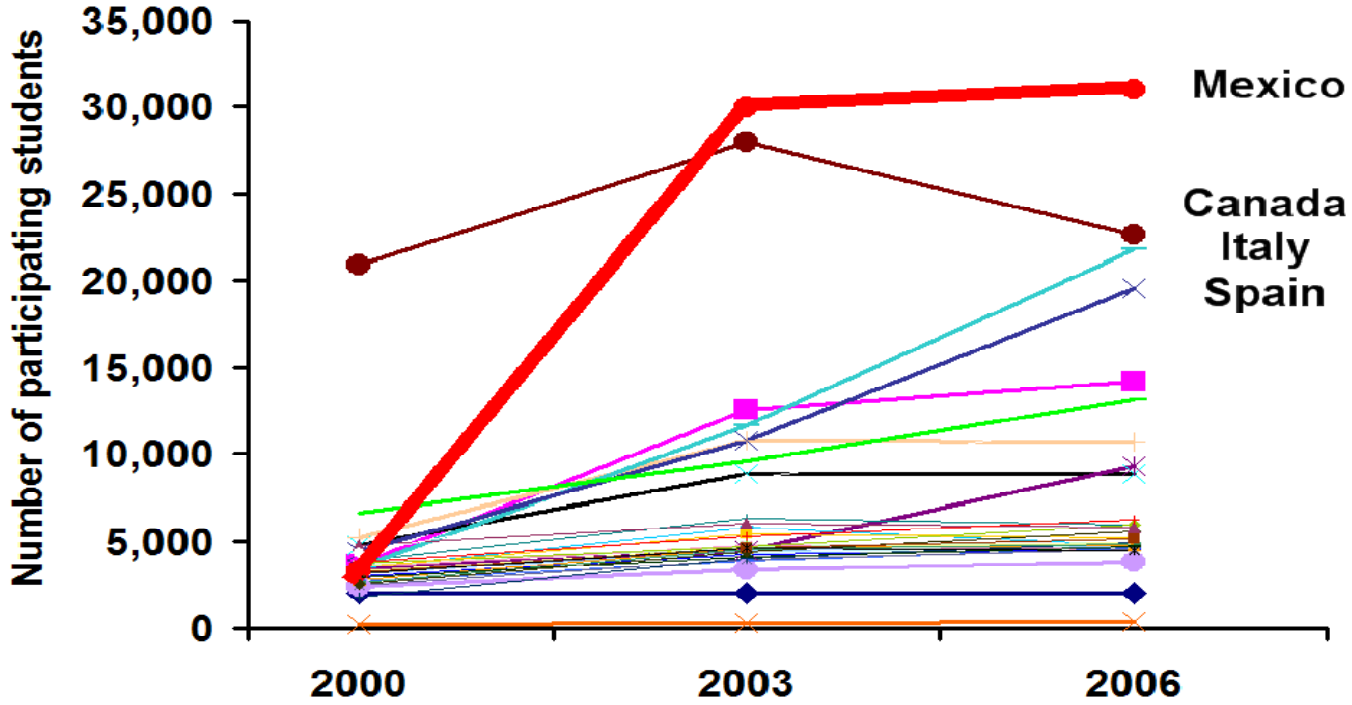
Better sample allocation

Option		State	Size	Frame				Sample		Rel. smp rates		deff	Effective sample
				# Schools		# Students		Schools	Students				
A	06	BAJA CALIFORNIA	Small	266	44%	1,915	5%	5	36	0.5	1.0	1.03	1,261
	05	BAJA CALIFORNIA	Medium	106	18%	2,652	7%	3	75	0.7	1.5		
	04	BAJA CALIFORNIA	Large	229	38%	31,053	87%	34	1,190	1.0	2.0		
		Total		601	100%	35,620	100%	42	1,301				
	96	ZACATECAS	Small	856	82%	3,313	21%	11	43	0.1	1.0	1.87	642
	95	ZACATECAS	Medium	77	7%	1,864	12%	6	145	0.8	6.1		
	94	ZACATECAS	Large	105	10%	10,768	68%	29	1,015	1.0	7.3		
		Total		1,038	100%	15,945	100%	46	1,203				
B	06	BAJA CALIFORNIA	Small					4	29			1.03	1,000
	05	BAJA CALIFORNIA	Medium					4	100				
	04	BAJA CALIFORNIA	Large					26	900				
		Total						34	1,029				
	96	ZACATECAS	Small					26	100			1.20	1,000
	95	ZACATECAS	Medium					6	149				
	94	ZACATECAS	Large					27	954				
		Total						59	1,203				

Option A	2,504	1,903
Option B	2,232	2,000
Total	272	-97

Mexico’s Participation

Sample Size

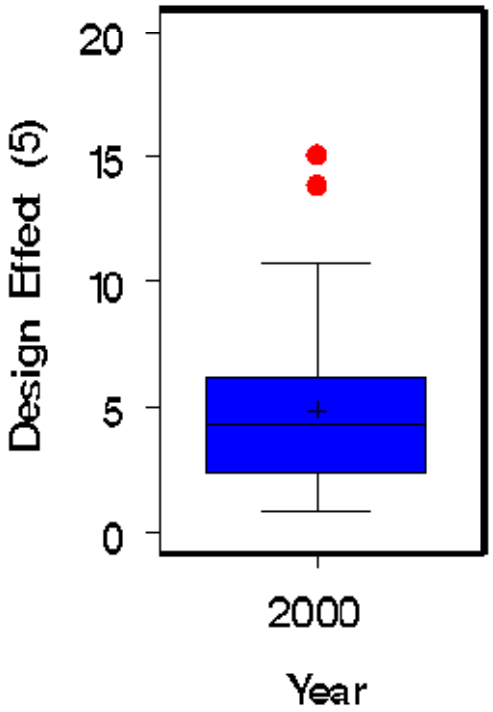


	Average Total Design Effect			
Year	Mexico	Canada	Italy	Spain
2000	5	6	4	5
2003	53	12	11	8
2006	34	12	11	14

Total Design Effect in 2000

- **DEFF (5)**

Number of Countries	34
Overall Mean	4.9
Mexico Mean	5.1



Quantile		Value
100%	Maximum	15.1
95%		13.9
90%		8.5
75%	Q3	6.2
50%	Median	4.2
25%	Q1	2.3
10%		1.8
5%		0.9
1%		0.8
0%	Minimum	0.8