

The Educational Characteristics of the Immigrants from Mexico

Alfredo Cuecuecha
ITAM-CIE
November 18



Motivation I

- Recent studies have found that:
 1. average years of education for Mexicans in Mexico < average years of education for Mexicans in the US
 2. migration probability is bigger for Mexicans with medium levels of education
 3. migration probability is also high for individuals with zero years of education



Motivation II

- Results are obtained using US and Mexico censuses
 - A. Advantages
 - 1. Representative
 - B. Disadvantages
 - 1. Endogenous sampling
inconsistent and inefficient results
 - 2. Undercounting of Mexicans in the US
inconsistent estimations



Objectives

- Obtain migration determinants
consistent and efficient
- Analyze the effect of undercounting
- Obtain estimations less affected by
undercounting



Endogenous sampling

1. Census samples require use of correct weights.

Example:

10 individuals in US

20 in Mexico

Is migration rate 30% or is it just that individuals in US are being sampled at higher rate?



Endogenous sampling: solution

- If sampling and aggregate migration rate are known consistent estimators are available (WESML)
- Such estimators depend on whether weights and aggregate rates are correct or not




Endogenous Sampling: solution II

- Imbens (1992) proposed a GMM estimator in which sampling rate and aggregate migration rate are parameters to estimate
- In principle, this solves both the endogenous sampling and the undercounting problem



Drawbacks

- GMM can have local minima or more than one global minima, consequently different starting points can generate different results
- Undercounted may be self selected, which generates inconsistent parameters



Local minima and multiple global minima

- Domínguez and Lobato (2004) have proposed an estimator to solve this problem (not implemented here)
- Alternative followed here is to attempt different starting points based on different years cross examining results against aggregate migration rates



Self selection in the undercounted

- An option is to supplement information with surveys about the characteristics of the undercounted
- Since illegal immigrants are likely to be over represented among the undercounted, a survey with their characteristics is key to solve the problem



Data

- Mexico: 1990 census (1% sample) and 2000 census (random sample from the stratified 9.1% sample)
- US: 1990 census (5% sample) and 2000 census (random sample from the 5% sample)
- Objective: two cross sections with equal number of individuals



Sample

- Mexico:
 1. Males
 2. 25 to 58 years old
 3. In the labor force
- EU:
 - 1 to 3

Only individuals that entered after age 17 to the US are included

	1990		2000	
	Mexico ^b	US ^d	Mexico ^c	US ^d
In Labor Force	114,771 (90.17%)	46,303 (92.32%)	1,424,894 (88.68%)	85,734 (75.76%)
Out of Labor Force	12,515 (9.83%)	3,850 (7.68%)	181,940 (11.32%)	27,424 (24.24%)

^aAll individuals in the US sample entered that country after their 17th birthday
^b1 % sample from the 1990 Mexico census
^cCensus sample of 2.2 million households and 9.1% individuals from the 2000 Mexico census
^d5% sample from the 1990 and 2000 US census

	1990			2000		
	MX ^b	US ^c		MX ^d	US ^e	
		All	Recent ^f		All	Recent ^f
Education	6.76 (5.08)	7.17 (4.71)	7.82 (4.93)	7.63 (4.86)	8.31 (4.53)	8.73 (4.42)
Age	37.55 (9.1)	37.34 (8.73)	32.73 (7.73)	37.80 (9.00)	37.62 (8.74)	33.05 (7.63)
Weight	100.76 (.8405)	20.69 (9.58)	20.59 (10.13)	10.19 (8.09)	20.77 (11.16)	21.30 (12.09)
N	114,771	46,303	10,964	114,771	46,303	11,640

^aAll individuals in the US sample entered after their 17th birthday
^b1 % sample from Mexico 1990 Census
^c5 % sample from US 1990 Census
^dRandom sample from the 9.1% sample of the 2000 Mexico census that matches the number of observations from b.
^eRandom sample from the 5% sample of the 2000 US census that matches the number of
^fDefined as individuals that entered during the five years previous to the census

Aggregate Data

- Sampling rates obtained from weights provided by census authorities
- Aggregate migration rate obtained with different sources

Table 1, Migration Rates, 1990-2000

	Mexicans in Mexico	Mexicans in the US	Mexicans in the US plus Undercounted	Migration Rate with no Undercounting	Migration Rate with Undercounting
1990	81.25*	4,30**	5,72	5,02%	6,57%
1995	91.06*	6,68**	8,88	6,83%	8,88%
2000	97.48*	9,18**	12,21	8,60%	11,13%

Source:

* INEGI, 2005; ** Chiquiar and Hanson, 2005. Undercounting rate of 33% as reported in Bean et al (2002).

Data

- MMP 93
 1. includes surveys in 93 rural communities and urban neighborhoods (1982- 2002)
 2. surveys in Mexico during winter, in summer surveys in the US
 3. Representative of communities that samples

Table 3, Mexican Males 25 to 58 years old* found in the MMP93

Survey rounds	Non Migrants		Migrants			
	1982-1992	1992-2002	1982-1992		1992-2002	
Education Category			All	Illegal	All	Illegal
No education	14.33	7.71	21.73	20.83	11.43	11.41
1 to 4 years	31.46	18.90	41.64	42.53	34.05	32.54
5 to 8 years	28.76	31.85	25.50	25.93	33.27	33.63
9 to 12 years	14.68	26.34	7.89	6.97	16.20	17.02
12 to 16 years	7.62	8.16	2.59	3.05	3.21	3.14
16+	3.16	7.04	.65	.69	1.84	2.26
N	1997	4457	1698	1018	2179	1463

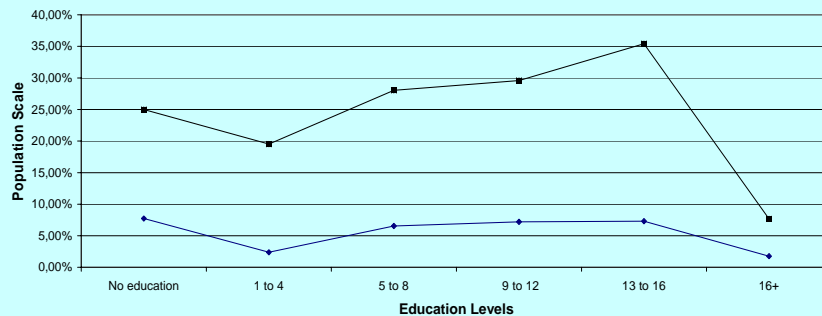
*Age for non-migrants at the moment of the survey. Age for migrants taken at the moment of their last trip. Includes only individuals that migrated to the US after age 17.

Added observations

Education Category	1990	2000
No education	3183	1743
1 to 4 years	6499	4972
5 to 8 years	3962	5139
9 to 12 years	1065	2601
12 to 16 years	466	480
16+	105	345
N	15280	15280

are census weights right?

Weighted Share of Mexicans in the US by Educational Attainment*, US and Mexico
Census 1990 and 2000, 25 to 58 years old males in the labor force**



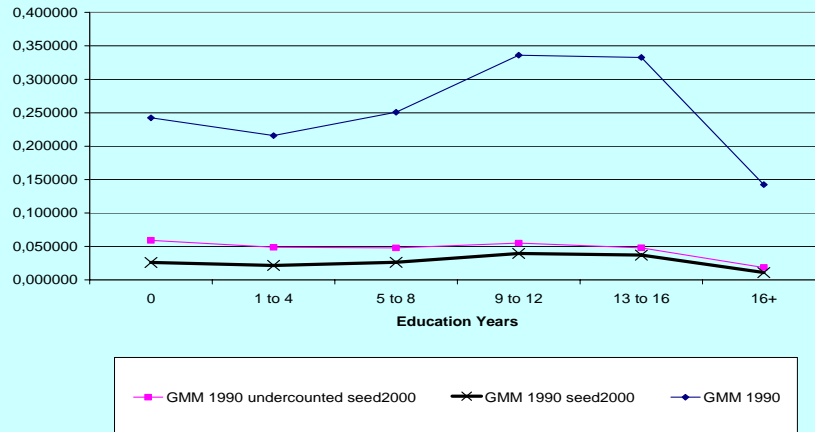
*Weights created according to Manski and Lerman (1977)

**All males in the US entered after their 17th birthday

■ Population 1990 ◆ Population 2000

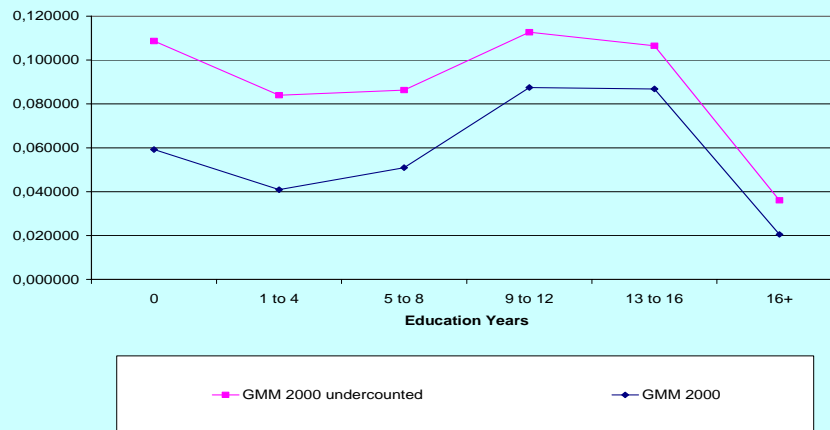
does GMM solves the problem?

Average Migration Probability by Education Categories in US Census, Males in the Labor Force, 1942 to 1965 Cohorts, 3rd degree Polynomial, 1990



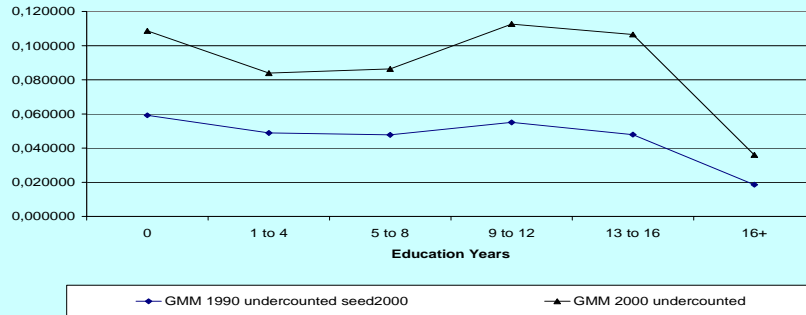
Does undercounting matters?

Average Migration Probability by Education Categories in US Census, Males in the Labor Force, 1942 to 1965 Cohorts, 3rd degree Polynomial, 2000



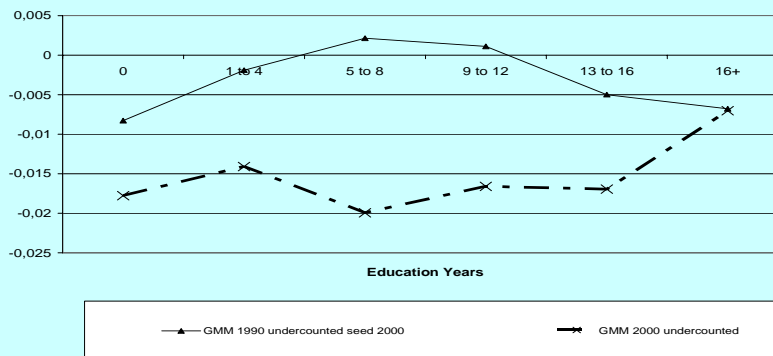
Migration rates 1990 and 2000

Average Migration Probability by Education Categories in US Census, Males in the Labor Force, 1942 to 1965 Cohorts, 3rd degree Polynomial, US and Mexico census



Marginal effects

Average Marginal Effect of Education by Education Categories in US Census, Males in the Labor Force, 1942 to 1965 Cohorts, 3rd degree Polynomial



What is behind the results?

Table 5
WESML and GMM Estimations, 25 to 58 Males in the Labor Force^a
US and Mexico Census 1990
3rd degree

	WESML	GMM		
		Seed 1990	Seed 2000	Seed 2000 ^c
N	161,074	161,074	161,074	176,354
Education	-0,0887974* (.0048829)	-0,08925* (.00002)	-0,0757941* (.00014)	-0,0704295* (.00035)
Education ²	0,0210058* (.0007545)	0,021282* (.00003)	.0173213* (.0013)	.0129795* (.000055)
Education ³	-0,0009474* (.0000308)	-0,000941* (.00000005)	-.0007763* (.00007)	-.0005748* (.000022)
Sampling Rate in US ^b	0.0765	0,0761* (.0004)	.4510* (.04302)	.440788* (.01256)
Predicted Migration At sample mean	.2509 (.0510)	.2816 (.0568)	.0307 (.0079)	.0532 (.0135)
Marginal Effect Of Education	.0007 (.0194)	.0019 (.0198)	-.0001 (.0032)	-.0015 (.0042)
Pseudo R ² / GMM objective function	.0151	.93802	.7517	.6498
Hausman** χ^2		.3468	.000001	nv

^aAll males in the US sample entered the US after age 17th.
^bFor the WESML estimation this variable is data used in estimation.
^cSample including the undercounted
*Significant at 1%
**For the test of no difference in GMM and WESML parameters. Butler (2000) modified version; does not reject null at 1%
nv: Modified Hausman test is nonvalid.

What is behind the results?

Table 6
WESML and GMM Estimations, 25 to 58 Males in the Labor Force^a
US and Mexico Census 2000

	WESML	GMM	
			Undercounted
N	161,074	161,074	176,354
Education	-.14345* (.0050553)	-.1433137 (.1034)	-.107582* (.0004)
Education ²	.02844* (.0007662)	.0284201 (.3644)	.019536* (.00003)
Education ³	-.0012086* (.0000312)	-.0012095 (.018963)	-.000828* (.00000001)
Sampling Rate in US ^b	.45109	.4505	.451099* (.00057)
Predicted Migration At sample mean	.0549 (.0028)	.0632 (.0767)	.0872 (.0204)
Marginal Effect Of Education	0.0001 (.0055)	-.0061 (.0050)	-.0165* (.0050)
Pseudo R ² / GMM objective function	.0179	.9743	.9987
Hausman** χ^2		nv	.0000001

^aAll males in the US sample entered the US after age 17th.
^bFor the WESML estimation this variable is data used in estimation.
*Significant at 1%
**For the test of no difference in GMM and WESML parameters. Butler (2000) modified version; does not reject null at 1%
nv: Modified Hausman test is nonvalid.



Final Remarks

- Tests show that the use of GMM or WESML is not as important as the use of the right weights
- These are not always easy to obtain with GMM due to local minima and multiple global minima
- Self selection in the undercounted can change results dramatically