Vocational Training for Disadvantaged Youth in Colombia:

A Long Term Follow up

Orazio Attanasio, University College London, IFS and NBER

Arlen Guarín, Carlos Medina, *Banco de la República*

Costas Meghir, Yale University, NBER and IFS

Online Appendix

A. Jóvenes en Acción

The Intervention.—The *Jóvenes en Acción* program was a training program for urban young unemployed that was implemented in the early 2000s in Colombia as part of a wider strategy called *Red de Apoyo Social*, aimed at providing a safety net for the poorest sectors of the population after the crisis that hit the country in the late 1990s.[1] It was initially funded with a USD$70 million loan from the Inter-American Development Bank, and was targeted to unemployed youths 18 to 25, who belonged to the poorest population classified in the two lowest levels of a score, called SISBEN, which is used in Colombia to target all welfare programs.[2] As DNP (2000b) highlights, unemployment rates of youths between 18 and 25 years of age of the ten largest metropolitan areas, living in the first and second lowest deciles of the income distribution, at the turn of the Millennium were 62.8 percent and 52.8 percent respectively. The program was implemented in the seven main cities of the country: Bogotá, Medellín, Cali, Barranquilla, Bucaramanga,

[1] The Colombian crisis of the late 1990s is described by Medina, Núñez and Tamayo (2013). The first program of the *Red de Apoyo Social* was created by means of the CONPES policy document Number 3075 of March 15, 2000 (DNP, 2000a), and was aimed at generating employment for the poorest, while program *Jóvenes en Acción* was simultaneously created with program *Familias en Acción*, by the CONPES policy document 3081 of June 28, 2000 (DNP, 2000b), with a budget for both programs of USD$426 million, USD$320 million of which were to be provided by The World Bank and the Inter-American Development Bank.

[2] SISBEN is the acronym in Spanish for Information System for Beneficiaries Selection, and it consists of six levels built with the quality of life SISBEN score, used in Colombia to target public subsidies. To apply to the program, individuals were additionally required to have a valid id, and if the applicant was a mother of children under seven years of age, she must present an official document to prove her maternity. Applicants that had previously taken training courses at SENA or any training institution were also eligible to *Jóvenes en Acción*.

Manizales and Cartagena. According to FIP and DNP (2001), by 2001, the program planned to enroll about 100,000 students in these cities between 2001 and 2003. According to AKM, however, it actually began to enroll them in 2002, and, by 2005, it had enrolled 80,000 students.

The goal of the program, which we describe in more detail in Appendix A, was to increase the employability of the young beneficiaries and provide them support for building what was their *project of life*. More specifically, its objectives were: (i) to develop the youths' occupational skills, so as to increase their likelihood of employment and improve their performance at work, (ii) to promote the private supply of relevant job training programs for poor youths, and (iii) to put bring closer productive youths form poorer backgrounds and training institutions to firms.

Jóvenes en Acción consisted of training courses designed and provided by private institutions, known as ECAP. Each course was expected to train about 30 unemployed youths selected amongst eligible applicants. The course had to have three main components: (i) classroom training; (ii) on-the-job training; and (iii) the youth's *project of life* (FIP and DNP, 2001). The aim was to develop occupational skills, social skills and broader career objectives. The program also included a small stipend of about USD$2.20 per day for trainees without children under seven years of age, and to about USD$3.00 per day for women with children under seven.[3]

The training institutions designing and offering the courses could be for profit or not for profit and they had to satisfy certain criteria, listed in Appendix A. AKM report that 43 percent of them were for profit. In 2005 there were 114 ECAPs offering 441 courses to 26,615 trainees, with their instructors teaching about 7.6 hours per day.[4]

[3] Transfers to women with children under seven year of age were not contingent on the number of their children under seven, and were paid weekly per day in which beneficiaries had attended their courses during their classes, and biweekly, after they completed their training, during their training period. See FIP and DNP (2001), DNP (2008, 2002), and AKM. Beneficiaries were also covered by personal accidents insurance, and by civil liability insurance, but not health insurance, as most of the beneficiaries were likely to be covered by the non-contributory health system that insures the poorest Colombians.

[4] 40 percent of the beneficiaries were from Bogotá, 18 percent from Medellín, 16 percent from Cali, 11 percent from Barranquilla, 7 percent from Bucaramanga, 5 percent from Cartagena, and 2 percent from Manizales. The total amount invested was US$22 Million (See *Ministerio de la Protección Social*, 2005)

A unique feature of the program was that the ECAPs were paid only a relatively small amount if a youth only completed the three-month classroom component. A substantial fraction of the overall payment, instead, was conditioned on the student completing a three-month apprenticeship with the participating firms in a timely fashion (see FIP and DNP, 2001). Moreover, the ECAPs would receive additional payments if the beneficiaries were hired by the firms that trained them. The ECAPs had to screen the candidates who would be eligible for training, which allowed them to identify those most likely to succeed. This defined the experimental population. This incentive scheme, which stressed the need to identify skills needed in the labor market, was one of the most innovative characteristics of the program, especially compared to the training programs operated since 1957 at SENA, Colombia's government institution providing training, which often did not relate to the demand for specific skills in the labor market (See Saavedra and Medina, 2013). To the government, the total cost of the program is the sum of the cost of the course, plus the maintenance transfers disbursed during the six months (See FIP, 2011 and DNP, 2001). The evaluation was based on the last cohort to be trained, namely in 2005.

Program details: The first component was expected to last between 280 and 350 hours and was focused both on the development of basic abilities for becoming employed (independent of the specific field), and the development of occupation specific skills. The former objective was pursued by providing the youths with basic social abilities and developing their soft skills: teaching them to be proactive, resourceful and open to feedback; improving their verbal and written communication skills; their analytic, deductive and daily work problems solving skills; by encouraging them to assimilate and understand their job's organizational environment; by developing teamwork skills, etc. The latter goal was pursued by providing training in the specific field of their interest, including the expertise in the use of equipment and tools, didactic material, and the procurement of services; products or services production, etc.

The second component consisted of three months of on-the-job training, and was about 480 hours long, conditional on the labor schedule of the specific firms

in which the youths were trained. The training institutions, ECAPs, when designing the training courses, they had to identify participating employers that would take the young trainees on an apprenticeship basis. The ECAPS also wrote a *training plan* to facilitate the completion in the firm of the training process that began in their classrooms. It also includes an assessment of the youths' performance in terms of their achievements, agreed upon by the firm, the ECAP, and the youth.

The third component, the *project of life*, pursued the youths' comprehensive human development, orienting them towards a positive visualization of their abilities, personal and work perspectives; providing them with tools for decision-making. It encouraged the youths to reflect on their work, their imminent insertion in the labor force, and its meaning with respect to their future labor market perspectives, helping them build their labor identity. This component took place all through the six months of the intervention (See FIP, 2011 and DNP, 2001, Annex 7).

B. The SISBEN survey

Virtually all social programs in Colombia are targeted through an index known as SISBEN. This index is constructed as a weighted average of a number of household level variables[5] collected by the municipalities using a survey. The survey in urban areas covers all those living in neighborhoods classified in the lowest three out of six socio-economic strata[6]; in rural areas the survey covers most (if not all) the population. In 2005, the SISBEN surveyed 32 million out of a total population of 43 million people. Thus, the SISBEN survey includes roughly, people in the 60 percent lowest percentile of the income distribution.[7]

[5] The formula to compute the index and the variables that enter the index are not publicly known. Moreover, the components of the SISBEN and their weight are periodically updated.

[6] The socioeconomic strata are based on the characteristics of the dwelling, and on the amenities of the neighborhoods, and do not consider the households' characteristics.

[7] The SISBEN survey was first collected in 1992 by all Colombian municipalities. Following the initial collection whenever any household wanted its information be updated it had to apply to the municipal Department of Planning, in charge of the local SISBEN administration. The SISBEN was updated the first time for the whole country between 2003 and 2005, and then between 2009 and 2010. Since the first time it was updated, its records became much more reliable and its score much harder to manipulate (See Bottia et al., 2012).

In our analysis we match the data from both the 2005 evaluation survey, and the entire universe of 2005 eligible applicants, with the SISBEN data, which provides us with some baseline information. We achieve a 90 percent match, independent of treatment and control status. We therefore construct two long-run follow up data sets with the matched applicants, and the characteristics of their households at the date the survey information was collected. Since this survey is a census of the poorest population, applicants that were not matched to it are assumed to belong to the higher part of the income distribution.

Table A1a. Personal Characteristics and Experimental Balance - Evaluation Sample

	All		Women		Men	
	Control mean	Treatment-control difference (p-value)	Control mean	Treatment-control difference (p-value)	Control mean	Treatment-control difference (p-value)
Employment	0.500	0.016 (>0.5)	0.456	0.015 (>0.5)	0.555	0.016 (>0.5)
Paid Employment	0.346	0.027 (>0.5)	0.326	0.019 (>0.5)	0.371	0.038 (>0.5)
Contract	0.085	0.003 (>0.5)	0.069	0.015 (>0.5)	0.104	-0.012 (>0.5)
Formal	0.085	0.014 (>0.5)	0.064	0.026 (>0.5)	0.111	-0.001 (>0.5)
Wage and salary earnings	99,881	-401 (>0.5)	83,866	1,831 (>0.5)	119,387	-3,159 (>0.5)
Self-Employment earnings	23,165	-1,687 (>0.5)	13,935	3,423 (>0.5)	34,407	-7,999 (>0.5)
Days worked per month	12.04	0.202 (>0.5)	10.795	0.322 (>0.5)	13.557	0.054 (>0.5)
Hours worked per month	24.864	1.342 (>0.5)	22.122	1.535 (>0.5)	28.204	1.103 (>0.5)
Education	9.997	0.255 (0.023)**	9.935	0.226 (0.308)	10.073	0.29 (0.201)
Age	21.229	-0.192 (>0.5)	21.39	-0.288 (0.387)	21.031	-0.074 (>0.5)
Married	0.202	-0.025 (>0.5)	0.261	-0.014 (>0.5)	0.129	-0.038 (>0.5)
Observations	3932		2125		1807	

Notes: The table reports the difference in each variable between the treatment and control groups, controlling for site-by-course fixed effects. The p-values were estimated taking into account that there were multiple hypotheses, using the Romano and Wolf (2005), and Romano, Shaikh and Wolf (2008), on each of the 11 baseline variables based on the bootstrap standard errors stratified by city, gender and treatment status.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A1b. Personal Characteristics and Experimental Balance - Entire Cohort

	All		Women		Men	
	Control mean	Treatment -control difference (p-value)	Control mean	Treatment -control difference (p-value)	Control mean	Treatment -control difference (p-value)
Low Socieconomic Stratum	0.957	-0.001	0.960	-0.001	0.951	-0.002
		(>0.5)		(>0.5)		(>0.5)
Living in house or apartment	0.864	0.005	0.854	0.006	0.885	0.005
		(>0.5)		(>0.5)		(>0.5)
Living at home without threats (avalanches, flood…)	0.918	0.003	0.916	0.002	0.920	0.006
		(>0.5)		(>0.5)		(>0.5)
Age in 2005	22.11	-0.163	22.16	-0.164	21.99	-0.160
		(0.003)***		(0.019)**		(0.340)
Homeownership	0.474	0.000	0.452	0.004	0.520	-0.009
		(>0.5)		(>0.5)		(>0.5)
Household size	5.555	-0.052	5.606	-0.082	5.451	0.022
		(>0.5)		(>0.5)		(>0.5)
Education of the head of the household	5.622	0.066	5.636	0.087	5.594	0.014
		(>0.5)		(>0.5)		(>0.5)
Age in 2005 of the head of the household	44.55	0.414	43.36	0.272	47.00	0.755
		(>0.5)		(>0.5)		(>0.5)
Number of children under 5 years old	0.719	-0.035	0.857	-0.035	0.438	-0.035
		(0.364)		(>0.5)		(>0.5)
Number of adults over 65 years old	0.144	0.007	0.133	0.003	0.167	0.017
		(>0.5)		(>0.5)		(>0.5)
Sisbén Score	11.0	0.185	11.0	0.202	11.2	0.142
		(>0.5)		(>0.5)		(>0.5)
Squared Sisbén Score	161.1	3.947	158.4	4.637	166.5	2.283
		(>0.5)		(>0.5)		(>0.5)
Applicant is the head of the household	0.088	-0.008	0.083	-0.002	0.100	-0.022
		(>0.5)		(>0.5)		(0.261)
Applicant is the Spouse/partner of the head	0.127	-0.008	0.185	-0.011	0.007	-0.001
		(>0.5)		(>0.5)		(>0.5)
Observations	31054		21649		9405	

Notes: The table reports the difference in each variable between the treatment and control groups, controlling for site-by-course fixed effects. The p-values were estimated taking into account that there were multiple hypotheses, using the Romano and Wolf (2005), and Romano, Shaikh and Wolf (2008), on each of the 14 baseline variables based on the bootstrap standard errors stratified by city, gender and treatment status.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A2 Descriptive statistics for outcome variables

	Women	**Men**	**Women**	**Men**
	Evaluation Sample, Jul/08-Dec/14		**Entire cohort (2010 only)**	
Formal sector	0.334	0.478	0.313	0.451
Months in formal sector	26.0	37.3	3.8	5.4
	26.8	*28.0*	*4.9*	*5.1*
Formal Income♦	237,283	368,171	207,135	311,500
	406,631	*529,853*	*373,719*	*426,831*
Pension contributions♦	34,634	56,000	29,605	47,163
	64,082	*84,413*	*58,824*	*68,091*
Health contributions♦	26,249	40,657	25,521	38,753
.	*46,231*	*61,022*	*46,375*	*53,036*
Contributions to *Cajas*♦	8,044	13,304	6,891	11,249
	15,561	*21,108*	*14,487*	*16,908*
Contributions to *SENA*♦	2,898	4,729	3,350	5,515
	6,627	*9,029*	*6,602*	*8,299*

Standard Deviations in italics. ♦RCOP$: Colombian pesos of December 2013 – Average per month. First two columns: Observations 2125 women and 1807 men. Last two columns: Observations 21649 women and 9405 men. *Cajas:* administration of family subsidies for low-wage employees with children. *SENA*: public vocational training program.

Table A3. Contributions to Social Security – Entire Cohort

	All		Women		Men	
	Control Means	Coefficient	Control Means	Coefficient	Control Means	Coefficient
	(1)	(2)	(3)	(4)	(5)	(6)
Pensions	31,803	4,190***	26,259	3,497***	43,119	5,863***
	(61563)	(678)	(57732)	(758)	(67319)	(1418)
		[0]		[0]		[0]
Health	27,229	3,319	23,036	2,869***	35,786	4,406***
	(48894)	(531)	(46399)	(597)	(52600)	(1098)
		[0]		[0]		[0]
Cajas	7,499	929***	6,143	738***	10,265	1,390***
	(15170)	(168)	(14150)	(187)	(16727)	(353)
		[0]		[0]		[0]
SENA	3,673	434***	3,007	327***	5,032	692***
	(7134)	(82)	(6466)	(91)	(8165)	(173)
		[0]		[0]		[0]
Obs.	126,072	372,648	84,612	259,788	41,460	112,860

Romano-Wolf P-values (in square brackets) for 4 hypothesis for overall sample and 8 when we split to men and women. Standard errors (in parentheses) clustered at the applicant level. ***Significant at the 1 percent level. Effects in COP$ of 2013. Year 2010. Observations: 12 months x individuals.

All regressions control for site by-course fixed effects and the following pretreatment characteristics: whether the applicant is in a low socio-economic stratum, whether living in house or apartment, whether living at home without threats , age in 2005, homeownership, household size, education of the head of the household, age in 2005 of the head of the household, number of children under 5 years old and number of adults over 65 years old in the household, sisbén score and squared sisbén score, whether the applicant is the head of the household and whether the applicant is the spouse/partner of the head of the household.