

Heterogeneous treatment effects in development policy evaluation: The case of the Mexican *Estrategia 100×100*

Curtis Huffman Espinosa^{a,b} • Brenda Valdez Meneses^a

Abstract

A recent evaluation of the Mexican interinstitutional coordination program *Estrategia 100×100*, published by Coneval in 2013, did not find many satisfactory results with respect to its target population. However, in Coneval's evaluation all statistical quantities of interest were computed at the group level only, thereby overlooking individual within-group heterogeneity. In this paper we estimate the heterogeneity of these treatment effects as a function of the level of investment per capita received through the program's actions. We provide evidence that the program was in fact effective whenever it was accompanied by the required investment.

Key words: program evaluation, average treatment effects, heterogeneity, propensity score matching.

JEL Classification: C21.

Resumen

Una evaluación reciente del programa mexicano de coordinación interinstitucional *Estrategia 100×100*, publicada por el Coneval en 2013, no encontró todos los resultados esperados en su población objetivo. En dicha evaluación, sin embargo, las estimaciones se llevaron a cabo sólo a nivel grupal, omitiendo analizar la posible heterogeneidad intra-grupo del impacto del programa. En este artículo analizamos esta heterogeneidad como una función de la inversión per cápita erogada. Las pruebas realizadas indican que el

Manuscript received on March 20th, 2014; accepted on October 23rd, 2015.

- a** The authors are respectively PhD student at El Colegio de México (Mexico), and Director of Program Evaluation for the Comisión Nacional de Uso Eficiente de Energía (Conuee, Mexico) [National Commission for the Efficient Use of Energy]. The authors would like to thank Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval) [National Council for the Evaluation of Social Development Policy] for letting us use their data. Furthermore, the authors would like to specially thank Thania de la Garza, Janet Zamudio and the anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the final version of this paper. The usual disclaimer applies.
- b** Corresponding author: <chuffman@colmex.mx>.

programa obtuvo de hecho los resultados esperados cuando se acompañó de la inversión requerida.

Palabras clave: evaluación de programas, efecto medio del tratamiento, heterogeneidad, emparejamiento, probabilidad de participación.

INTRODUCTION

There is a growing body of work in the literature concerning poverty and inequality that recognizes that complex social problems, such as the reduction of regional inequalities, have multiple causes and therefore require the design and implementation of integral public policies that involve the participation of several governmental, social and private institutions which need to share information, objectives, goals and resources (Grindle, 2010; Ordaz Ocampo, 2012; Brown and Tandon, 1992; Lustig, 2001). According to Brown and Kalegoankar (2000) and Gray (1989), it is almost impossible to efficiently counteract poverty and regional inequalities by improving coverage, efficiency and sustainability of basic public delivery services without encouraging the participation of multiple actors.

It is widely recognised that integral public policies that require the coordination of several government and private institutions are key elements in both poverty and regional inequality reduction strategies. Collaboration, cooperation and coordination among different levels of government, the private sector and the civil society are aimed at overcoming obstacles such as scarcity of government resources, different government capacities and priorities, and competing demands among the different sectors involved when governments try to address regional inequalities and poverty issues.

However, this interinstitutional coordination among the different actors involved has proven to be more complex in practice than it might seem. Efforts made in the past have shown that political will by itself is not enough for coordination strategies to work. Policy designers have to develop mechanisms and rules that suit the interests of multiple diverse institutions (Grindle, 2010) through differentiated social policies according to each local and regional reality.

For this reason, the Mexican government has been implementing integral policies since the seventies in order to address poverty and regional inequalities based on interministerial coordination. For instance, the last administration implemented the umbrella program *Estrategia 100×100* (E100×100) that sought to increase social and economic development in the 125 municipalities with the

lowest Human Development Index (HDI) by way of 300 different actions carried out through the coordination of several government ministries and institutions. Likewise, the current federal administration has just launched the *Cruzada Nacional Contra el Hambre* [National Crusade Against Hunger] that has targeted the population that lives in extreme poverty with food deficiencies.

Despite advances in the role of social policy in economic development—the creation of the *Consejo Nacional de Evaluación de la Política de Desarrollo Social* (Coneval) [National Council for the Evaluation of Social Development Policy], for instance—the interinstitutional coordination strategy launched by the Mexican government in recent years seems to not have had very satisfactory results in the vulnerable population they have targeted. A recent impact evaluation of E100×100 published by Coneval (2013) only found a positive impact in the reduction of the percentage of population without health insurance and the percentage of households without concrete flooring.

These results posed severe doubts concerning both the effectiveness of the E100×100 program and the cooperative efforts of governments at different levels for alleviating poverty and reducing regional inequalities. However, this evaluation was based on an Intent to Treat Model (ITM) which means that the original assignment of the municipalities to E100×100 was used to carry out the analysis instead of the effective reception of the treatment, mainly due to the difficulty of measuring the actual interinstitutional coordination. Hence, as Outhwaite and Turner (2008) suggested, under the ITM, finding a non-statistically significant difference between the treatment and the control group cannot be interpreted as if receiving or not the treatment produces equivalent results. It could have been that the treatment spilled out to some of the control units—as argued by Coneval in their evaluation. In Coneval's evaluation, all statistical quantities of interest were computed only at the group level, overlooking individual within-group level variations or heterogeneity which, according to the data provided by *Secretaría de Desarrollo Social* (Sedesol) [Ministry of Social Development], exists in terms of the investment per capita received through the program's specific actions across treated municipalities. It is important to mention that the investment data was available only for municipalities that were part of E100×100, so the possibility of using investment as the treatment variable was not possible.

The hypothesis of this paper is that E100×100's intervention is best captured not only by the treatment status of the municipalities, but also through

the different levels of investment per capita among the treated municipalities. Therefore, the results varied widely, depending upon whether E100×100's intervention included the financial resources dearly needed in the treated municipalities, diluting any treatment effect estimated using only the assignment to the program as treatment variable. In this paper, we put to empirical test this hypothesis assessing the heterogeneity in previously estimated treatment effects as a function of the level of investment per capita. Unfortunately, whether the observed variability in investment per capita speaks of the degree of coordination actually achieved by the officials involved is a whole different matter which we cannot assert with the available data.

There are, however, two reasons why we see the investment per capita as part of E100×100's intervention, irrespective of the terms in which this might have happened. Firstly, all the investment we have taken into account in our analysis ran through the E100×100. That is, we have only used the investment for which the coordination effort can be credited. Secondly, even though the investment was prominently federal (87%), this does not mean it required no involvement of the E100×100. All federal social programs that worked within the framework of E100×100, and through which the investments were made, had as part of their target population all of the 125 municipalities on which the E100×100 was focused. This left in sight no other obvious culprit for the observed differences in investment per capita but the inner workings of E100×100, especially among paired municipalities.

A word of caution is warranted though. Admittedly, exactly how correlated is this investment data with an actual measure of E100×100's effectiveness is a pending task in the literature, and our results should be read accordingly. Ultimately, the readers can make their own conclusions based on the heterogeneity analysis presented here.

This analysis was performed following the practical approach proposed by Xie, Brand, and Jann (2012)¹ based on observational data using Propensity Score Matching (PSM). Hence, the empirical strategy consisted of firstly, the replication of the Propensity Score (PS) estimates carried out in Coneval's evaluation using the same set of observed pre-treatment covariates for each municipality with the same sample consisting of the 2 456 municipalities throughout the country. Secondly, a multiple matching based on the PS estimated was used to transform

1 See also Brand and Thomas (2013) and Xie (2013).

the data into treatment-control comparisons replicating the estimation of Average Treatment Effects (ATT). Thirdly, the observed difference in pairs between treated municipalities and the group mean of the multiple matched controls (the matched difference for the ATT) was used to estimate the heterogeneous treatment effects as a function of the relevant investment per capita data by fitting a linear parametric model to evaluate trends across the municipalities.

The paper is organised as follows: Section 2 succinctly describes the coordination efforts implemented by the Mexican government up till the E100×100 umbrella program was launched. Section 3 reviews the impact evaluation published by Coneval. Section 4 describes the data and the methodology employed to assess the heterogeneity of the treatment effects of E100×100 as a function of the different levels of investment per capita received through the program's specific actions. Section 5 presents the key results and section 6 discusses and concludes these findings.

PREVIOUS COORDINATION EFFORTS UP TO THE E100×100 PROGRAM

Since the seventies, the Mexican government has been trying to alleviate complex social problems with the implementation and design of integral strategies based on interinstitutional coordination. In this sense, the idea of the necessity of integral public policies in Mexico is not new. However, as Ordaz Ocampo (2102) argues, the existence of power asymmetries has led to results that were less favourable than expected.

Probably the first attempt of this kind of programs occurred in 1973 with the implementation of the Programa de Inversión de Desarrollo Rural (PIDER) [Public Investment Program for Rural Development]. Its main objective was the promotion of integral rural development through the participation of the three levels of government. However, the integral approach and its transversal nature surpassed the public administration's capacity and ultimately generated actions that were not a product of this desired coordination among different government institutions (Herrera, 2007).

By 1977, the Coordinación General del Plan Nacional de Zonas Deprimidas y Grupos Marginados (COPLAMAR) [General Coordination for the National Plan of Depressed Zones and Marginal Groups], was created with the explicit objective of reducing levels of exclusion of farmers, migrants and indigenous

people, among others. According to Ordaz Ocampo (2012), both PIDER and COPLAMAR were undermined by the lack of public officials skilled in decision-making and with broad authority to determine the appropriate actions for local circumstances.

In 1994, the Congress declared as *Zonas de Atención Prioritaria* (ZAP) [Priority Attention Areas], the most vulnerable geographic areas on which development policies should focus their attention. This was part of the efforts for improving the design of public policies aimed at enhancing the integral development of the municipalities with the highest levels of social exclusion and poverty.

In 2000, the new administration launched the *Estrategia Microregiones* program [Micro-regions Strategy]. This program was meant to improve the wellbeing of the most disadvantaged part of the population who had been excluded from the provision of basic infrastructure and services due to their geographical dispersion and isolation, the low capacities of local governments. This involved the participation of multiple government officials, the private sector and the community, but in practice, the support came almost entirely from the highest Federal level mainly due to the change of the political party in power and the coexistence of different ideological views (ITESM, 2007).

With the same spirit of tackling regional inequalities and poverty, in 2006 the Mexican government launched yet another interinstitutional coordination program called *Estrategia 100×100* (E100×100), having as its main objective the increase of social and economic development in the 125 municipalities with the lowest HDI and highest levels of social exclusion.

The E100×100 umbrella program had two specific objectives: 1) To increase the income of the population who lived in these most disadvantageous 125 municipalities through actions oriented to increase productivity and employment opportunities and 2) To increase their quality of life through improvements in both the access and quality to health and education services as well as the housing and infrastructure conditions.

Through over 300 specific interventions carried out by several Government Ministries, the program was aimed at improving the wellbeing of the most vulnerable part of the population focusing on education, health, income, housing, infrastructure and environment. These interventions were grouped in what was then called “strategic actions” as shown in Table 1.

The E100×100 program was a coordination scheme meant to link and complement the different actions carried out by several government branches

through approximately 70 different development programs, all of them aimed at confronting regional inequalities and reducing poverty through the coordination of 14 Ministries and agencies of the Federal Government, 7 State Governments and 125 Municipality Governments over the 6 different aspects that were meant to improve the wellbeing of the population. Each of these 6 aspects of the E100×100 had an inter-ministry thematic working group within the Comisión Intersecretarial de Desarrollo Social (CIDS) [Interministry Commission for Social Development], which was in charge of establishing the conceptual and operational basis for the coordination both at the central and state level. The Technical Secretary presiding over these thematic working groups was the Secretaría de Desarrollo Social (Sedesol) [Minister of Social Development] with the goal of ensuring coordination and communication among them. Coordinating each thematic group was the responsibility of the head of the respective Ministry who was responsible for monitoring all investment plans and defining both the specific performance indicators and the targets to be achieved.

Table 1
Aspects and actions of the E100×100 program

<i>Aspect</i>	<i>Strategic actions</i>
Education	Shelters, literacy, scholarships, schools, educative packages and others
Infrastructure	Potable water, rural roads, highways, digital connectivity, urban development, electrification, bridges and sanitation
Income	Development agencies, training, saving and financing, temporal employment, support programs and productive projects
Environment	Environmental courses, environmental support projects, environmental projects, recycle, reforestation and ecological reserves
Health	Seniors, energetic support, food assistance, hospitals, clinics and mobile units, popular insurance and health services
Housing	Stoves, walls, housing packages, cement flooring, extra room, toilet services, ceilings, property titles and basic housing unit

Source: *Estrategia 100×100*. Available at: <<http://www.estrategia100x100.gob.mx>>.

In summation, the E100×100 was not a program of the Federal Government in the sense of having its own budget, but it was a scheme for the coordination of joint actions by different federal programs to promote the economic and social development. The E100×100 did not “plan” as much as it “manage”.

According to data provided by Sedesol for the period of 2007-2011, approximately 39 544 million pesos (some 2 950 million usd of the time) were spent

as part of the actions carried out by E100×100 from which 87% represented federal investment, 10% was state investment and only 3% represented municipal investment. From the grand total, the Sedesol contributed with the largest investment representing 45.6% of the total, followed by the Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CDI) [National Commission for the Development of Indigenous Populations] with 17.3%, while the Ministries of Education and Health only contributed 3.4% and 2.5% respectively.

PREVIOUS IMPACT EVALUATION OF E100×100

In order to identify the effects of E100×100 regarding one of its objectives—the improvement of living conditions—in 2012 Coneval carried out an intention-to-treat evaluation of the program by making use of public information of the 2,456 municipalities in the country, mainly from the *Censo de Población y Vivienda* [Population and Housing Census] from 2000 and 2010 and the *Conteo de Población y Vivienda* [Population and Housing Count] from 2005, drawing on this last one for baseline covariates.²

The causal treatment effects of the E100×100 program—under the assumption of treatment homogeneity—were estimated for several performance indicators including components of the HDI³, the Índice de Desarrollo Social (IRS)⁴ [Social Gap Index], and the Índice de Marginación (IM)⁵ [Marginalization Index]. The identification strategy to estimate the treatment effects was selection on observable factors, making use of PSM and Regression Discontinuity Design. Additionally, PSM was combined with Differences in Differences in order to control for unobservable characteristics that remained constant over time but that could have influenced selection into treatment and potential outcomes.⁶

Table 2 reproduces the main diagnostics of the different PSM algorithms employed in Coneval's evaluation.⁷

2 See Instituto Nacional de Estadística y Geografía. Available at: <www.inegi.org.mx>.

3 Provided by the PNUD (2000; 2005).

4 Provided by Coneval (2005).

5 Provided by CONAPO (2005).

6 For an introduction to these statistical methods see Gertler *et al.* (2011) and Bernal and Peña (2011), for a more specialized text see Guo and Fraser (2014), and Pan and Bai (2015).

7 For reasons of space we do not reproduce here the RDD's results, the details of which can be found in <http://www.coneval.gob.mx/Informes/Evaluacion/Impacto/Evaluacion_de_impacto_de_la_Estrategia_100x100.pdf>.

Several points are worth mention. It is important to notice that in Coneval's evaluation no good matching was possible for the 125 treated municipalities. With the best matching algorithm, radius caliper of 0.05 according to Coneval's evaluation, 52 treatment municipalities were discarded, that is, the treatment effect for these municipalities could not have been estimated because it was not possible to find them a good match, *i.e.*, the closest control municipality for each of them had a propensity score "farther away" than 0.05. Note that even though the inverse-probability weighting estimator (IPW) made use of most of the data, Table 2 shows that the *p*-values of the likelihood-ratio test of the joint insignificance of all the regressors, before and after matching, were both zero, and that it exhibited the second larger value for the median of the standardized bias.

Table 2
Diagnostic tests before and after matching

<i>Algorithm</i>	<i>Median bias</i>		<i>p>chi2</i>		<i>Observations after matching</i>	
	Before matching	After matching	Before matching	After matching	Treatment	Control
Nearest neighbor without replacement	156.1	18.15	0	0	74	74
Nearest neighbor with replacement	156.1	14.85	0	0.001	74	24
Nearest neighbor with replacement (5 neighbors)	156.1	12.67	0	0.087	74	53
Kernel (0.05)	156.1	23.42	0	0.263	73	152
Radius caliper 0.025	156.1	20.7	0	0.346	56	91
Radius caliper 0.05	156.1	11.1	0	0.168	73	156
Inverse-probability weighting (IPW)	156.1	20.40	0	0	119	1 632

Note: *p*-values correspond to the likelihood-ratio test of the joint insignificance of all the regressors before and after matching.

Source: Coneval (2013) and own calculations.

For completeness, Tables 3 and 4 reproduce, in levels and differences respectively and for every PSM algorithm, the estimated effect of the E100×100 program on the outcome variables analyzed in Coneval's evaluation.

Table 3
ART in levels of 2010

Outcome variables (2010)	Radius caliper			Nearest neighbor			Kernel	ipw
	0.025	0.05	Without rep	With rep	With 5 neighbors	Kernel		
% of illiterate population of 15 years or older	2.501 (2.253)	2.040 (1.990)	3.407** (1.33)	2.181 (2.203)	2.074 (1.81)	2.408 (1.990)	5.645*** (2.121)	
Rate of scholar attendance 6 to 24 years old	-1.123 (2.237)	-1.226 (1.968)	-0.241 (0.991)	-1.061 (2.352)	0.146 (1.85)	-1.412 (2.022)	0.0927 (2.773)	
Rate of scholar absence 6 to 14 years old	1.819 (1.396)	2.036* (1.152)	0.969 (0.717)	2.165 (1.327)	1.347 (1.061)	2.212* (1.182)	2.528* (1.391)	
Rate of incomplete elementary school	1.740 (2.015)	2.208 (1.883)	3.126** (1.195)	2.891 (2.069)	1.402 (1.79)	2.61 (1.92)	2.956 (2.318)	
% of population without access to health services	-26.6*** (8.105)	-22.4** (8.751)	-12.47** (5.242)	-26.65*** (9.204)	-19.39*** (8.37)	-23.74** (8.749)	-23.05** (10.38)	
% of households without cement flooring	2.469 (4.102)	2.094 (3.766)	6.009** (2.292)	3.295 (4.394)	1.94 (3.474)	2.202 (3.805)	-3.060 (5.031)	
Average number of persons per room	0.131* (0.0712)	0.115* (0.0674)	0.177*** (0.0578)	0.158* (0.081)	0.864 (0.067)	0.118* (0.07)	0.160*** (0.0521)	
% of households with neither drainage nor private toilet	2.048 (4.875)	2.400 (3.982)	1.08 (2.524)	3.035 (4.704)	0.745 (3.393)	3.178 (4.11)	0.124 (5.427)	
% of households without tap water	-7.732 (7.768)	-8.518 (7.527)	-1.675 (4.202)	-9.432 (9.106)	-3.373 (6.563)	-7.801 (7.606)	-18.34 (11.18)	
% of households without drainage system	3.367 (8.909)	2.700 (8.237)	8.352 (5.243)	3.197 (8.438)	4.934 (7.266)	2.562 (8.326)	-0.485 (8.862)	
% of households without electricity	1.955 (2.942)	2.375 (2.844)	-0.123 (1.935)	2.774 (3.258)	2.132 (2.251)	2.466 (2.849)	-2.033 (4.790)	
% of households without laundry	5.158 (3.618)	3.087 (3.502)	3.969** (1.547)	4.542 (4.281)	4.318 (2.733)	3.022 (3.503)	12.04 (7.901)	
% of households without refrigerator	4.923 (4.138)	5.724 (3.772)	8.146*** (2.791)	8.224* (4.725)	4.136 (3.371)	5.325 (3.71)	11.51** (5.760)	
% of population aged +15 without complete elementary school	1.850 (2.309)	2.235 (2.141)	3.152** (1.285)	2.945 (2.452)	1.388 (2.045)	2.728 (2.189)	3.091 (2.799)	

Poverty	-0.355 (0.880)	-0.534 (0.879)	-1.077 (0.86)	-0.353 (0.929)	-0.624 (0.795)	-0.503 (0.891)	-0.110 (0.603)
Extreme poverty 2010	-0.578 (2.311)	-0.265 (2.385)	2.909 (1.992)	0.314 (2.68)	-0.847 (2.246)	-0.267 (2.42)	2.228 (2.783)
Moderate poverty 2010	0.192 (1.883)	-0.290 (1.849)	-1.832 (1.399)	-0.692 (2.073)	0.198 (1.735)	-0.26 (1.861)	-2.354 (2.541)
Vulnerable population by social deprivation 2010	0.464 (0.814)	0.665 (0.785)	-0.747 (0.77)	0.455 (0.861)	0.695 (0.738)	0.623 (0.801)	0.269 (0.576)
Vulnerable population by income 2010	-0.0372 (0.0621)	-0.0531 (0.0749)	-0.0797* (0.045)	-0.0432 (0.065)	-0.0343 (0.049)	-0.05 (0.074)	-0.0954* (0.0565)
Not poor neither vulnerable population 2010	-0.0781 (0.126)	-0.0818 (0.147)	-0.250** (0.123)	-0.0568 (0.121)	-0.0392 (0.092)	-0.074 (0.146)	-0.0719 (0.107)
Deprived in the access to social security 2010	0.257 (1.002)	0.111 (1.055)	0.912 (0.586)	-0.0311 (1.082)	-0.127 (0.689)	0.0064 (1.05)	1.965 (1.803)
Deprived in the quality and spaces of housing 2010	7.168 (5.029)	7.285 (4.654)	11.07*** (2.784)	9.053 (5.513)	6.632 (4.257)	7.449 (4.754)	-1.166 (7.427)
Deprived in the access to social basic housing services	2.497 (2.895)	2.015 (2.669)	2.691 (2.161)	3.81 (3.247)	3.271 (2.68)	2.401 (2.742)	2.263 (3.300)
Deprived in the access to food	0.217 (3.228)	-1.323 (2.822)	1.2 (2.318)	-0.238 (3.148)	-2.459 (2.728)	-0.942 (2.882)	1.635 (2.747)
Population with at least one social deprivation	0.116 (0.178)	0.133 (0.217)	0.335** (0.159)	0.0986 (0.176)	0.0735 (0.129)	0.122 (0.215)	0.162 (0.150)
Population with three or more social deprivations	-1.849 (3.099)	-1.430 (3.301)	1.077 (1.914)	-1.534 (3.458)	-2.025 (2.623)	-1.543 (3.309)	1.872 (4.828)
Population with income below the welfare line	-0.397 (0.860)	-0.586 (0.846)	0.984 (0.835)	-0.4 (0.907)	-0.66 (0.778)	-0.552 (0.859)	-0.202 (0.603)
Population with income below the minimum welfare line	0.430 (1.765)	0.295 (1.764)	3.269** (1.554)	0.847 (1.99)	0.932 (1.617)	0.339 (1.806)	1.415 (1.552)
Observations							
Treatment	56	73	74	74	74	73	119
Control	91	156	74	24	53	152	1 632

Source: Coneval (2013) and own calculations.

Table 4
ARR in Differences in Differences

Outcome variables	Radius caliper			Nearest neighbor			Kernel	IPW
	RC 0.25	RC 0.05	Without rep	With rep	With 5 neighbors	Kernel		
% of illiterate population of 15 years or older	2.362 (2.18)	2.561 (1.9)	0.797 (1.519)	2.397 (1.906)	2.607 (1.958)	2.701 (1.879)	0.983 (2.075)	
Rate of scholar attendance 6 to 24 years old	0.908 (2.3)	-0.116 (2.6)	0.47 (1.264)	0.157 (2.572)	-0.627 (2.134)	0.169 (2.558)	3.899 (5.143)	
Rate of scholar absence 6 to 14 years old	-0.386 (1.59)	-0.24 (1.39)	-0.284 (1.245)	-0.655 (1.622)	-0.387 (1.524)	-0.417 (1.423)	-0.00691 (1.118)	
% of population without access to health services	-30.37** (13.32)	-30.37** (12.85)	-14.14 (10.21)	-32.75** (14.87)	-33.12** (12.67)	-30.64** (12.78)	-19.62** (9.613)	
% of households without cement flooring	-12.8 (8.79)	-14.56* (8.8)	-14.85** (6.152)	-12.69 (9.219)	-16.50** (8.155)	-13.64 (8.86)	-14.49** (6.933)	
Average number of persons per room	0.161 (0.27)	0.224 (0.221)	0.224 (0.15)	0.161 (0.268)	0.272 (0.203)	0.178 (0.234)	0.254 (0.234)	
% of households with neither drainage nor private toilet	7.368 (5.45)	4.73 (4.12)	0.241 (3.199)	6.296 (5.956)	3.203 (5.233)	5.694 (4.311)	5.339 (4.534)	
% of households without tap water	-0.561 (8.63)	-4.64 (7.72)	-5.575 (6.123)	-0.296 (7.689)	-4.668 (7.077)	-3.625 (7.745)	1.263 (7.524)	
% of households without drainage system	2.52 (8.95)	3.17 (7.65)	-1.285 (6.096)	5.522 (8.881)	4.433 (7.333)	3.211 (7.76)	5.169 (5.530)	
% of households without electricity	-2.61 (4.24)	-1.46 (3.72)	1.112 (2.907)	-2.407 (3.981)	-0.337 (3.141)	-1.682 (3.759)	-3.048 (3.230)	
% of households without laundry	-2.378 (2.017)	-2.00 (1.79)	-0.214 (1.021)	-1.752 (2.115)	-0.675 (1.375)	-2.164 (1.815)	-4.426 (3.413)	
% of households without refrigerator	-4.17 (3.4)	-3.98 (2.97)	-2.904 (1.878)	-3.1 (2.925)	-2.295 (2.523)	-4.081 (2.984)	-7.730 (5.706)	

Note: The differences taken correspond to the years 2010-2005-(2005-2000). Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: Coneval (2013) and own calculations.

From the 29 outcome variables analyzed in this evaluation, according to the matching algorithm favored by Coneval, statistically significant results were found only in the reduction of the percentage of population without access to health services and in the reduction of the percentage of households without cement flooring. However, this last result was not robust for all the different samples and specifications employed. The improvement in conditions regarding the lack of access to health services was 30.4%. A figure that represents a reduction in the performance indicator of 47.3% with respect to the control group (Coneval, 2013).

However, it is important to recall that this evaluation was based on the ITM so if the results did not show significant differences between the treatment and the control municipalities in the relevant outcome variables over which the impact was measured, this should not be interpreted as if the E100×100 program did not have the desired impact on their beneficiaries. As mentioned before, under the ITM, it was not possible to determine whether some of the treated municipalities ended up being untreated or if some of the control municipalities actually received the treatment due to spillover effects. In other words, all we know is that municipalities in the sample ended up being treated alike —*i.e.* missing or receiving the treatment— and therefore responding alike (Outhwaite and Turner 2008) on the average.

Accordingly, the conclusions of the evaluation performed by Coneval state that it could have been the case that the actions of the different federal social programs coordinated by the E100×100 umbrella program did not focus only on the 125 most vulnerable municipalities, but covered all municipalities with high IRS, the control group among them. In summary, there was no evidence to state that there was a singular prioritization of these 125 municipalities in contrast with other poor municipalities of the country. It merits attention the fact that the only outcomes in which the evaluation found the expected results were associated to two highly centralized subprograms: *Seguro Popular* [Popular Insurance] and *Piso Firme* [Concrete Flooring], presumably two subprograms that didn't require much coordination.

In Coneval's evaluation, all statistical quantities of interest were computed only at the group level —under the assumption of treatment homogeneity—, overlooking within-group heterogeneity which existed in terms of the investment per capita received through the program's specific actions across different municipalities. Heterogeneous treatment effects are seldom studied empirically

in the evaluation of development programs, regardless of their important implications for social policy. On the one hand, if policy makers understand patterns of treatment effect heterogeneity, they can more effectively concentrate their efforts where they are more sorely needed. On the other hand, the study of treatment effect heterogeneity can also yield important insights about how scarce social resources are distributed in an unequal society (Xie, Brand, and Jann, 2012; Manski, 2007).

In this paper we followed Xie, Brand, and Jann (2012) in their approach to studying treatment effect heterogeneity that builds on the same framework for estimating the causal effects. That is, under the ignorability or selection on observables assumption. The ignorability assumption allows the researcher to explore empirical patterns of treatment effect heterogeneity as a function of other variables such as the propensity score itself and in our case of investment per capita disaggregated by specific action of the program, which in turn it was taken as differences in the actual treatment received by the municipalities. This strategy was followed in view of the lack of any additional information regarding the actual coordination achieved by program.

DATA

For the replication of the results estimated in Coneval's evaluation, we used the same data on socioeconomic conditions, social gaps, levels of marginalization and poverty from the 2 456 municipalities of the country contained in the *Censo de Población y Vivienda* of 2000 and 2010, the *Conteo de Población y Vivienda* of 2005, the municipal HDI, the IRS and the IM.

The 125 treated municipalities of the E100×100 program comprised the 100 municipalities with the lowest HDI in the year of 2000 and the 25 municipalities, identified by Coneval, both with a high level of social exclusion and more than 60 per cent of their population living under food poverty. Whereas the pool of control municipalities consisted of the rest of the municipalities of the country. That is, 2 331 municipalities comprised the pool from which the comparison group were matched.

Additionally, for the estimation of the heterogeneity in treatment effects, this paper used information provided by the Sedesol regarding the level of investment by Strategic Actions for each of the 125 treated municipalities from 2007 to 2011. It is important to recall that this investment data was not avail-

able for the 2 331 non-treated municipalities of the sample. Note also that the investment data we used in our estimates was produced as a byproduct of the E100×100 and no similar data exists for any other municipality apart from the 125 that comprised its target population. By no means this is to say that none of the other municipalities excluded from the E100×100 received any investment, or that they were not part of the target population of other, or even the same, social programs. The social programs themselves did not constitute the E100×100 for it was a coordination scheme.

This is why it was impossible to use the investment variable as the treatment. Instead, the investment per capita by Strategic Action in each of the 125 treated municipalities was taken as differences in the intervention that actually took place.

According to the yearly investment data provided by Sedesol, the global investment per capita by Aspect at the municipal level has been unequal among treated municipalities for the period 2007-2011. For instance, the municipality of Mezquital in the northern State of Durango received 8.7 more resources per capita than San Simón Zahuatlán in the south state of Oaxaca for the period of 2007-2011. On average the former municipality received 3.1 more resources than the rest of the municipalities that were part of E100×100. In this paper we exploited this variability to get a feeling of the effect that different versions of E100×100's intervention had on the outcome variables employed in Coneval's analysis.

Regarding the quality of the investment data, it is important to keep in mind that we used it as it was provided by Sedesol: with minimum disaggregation of the specific institution financially responsible for the spending or the specific kind of service delivered. These were aggregate figures and should not be interpreted as an exhaustive account of the money spent in the municipalities that comprised the E100×100 target population in that period. However, it was not our intention to use the investment data to estimate the true cost of observing a desired effect on the most vulnerable territories in the country. Rather, in this paper we used the variability in the data as an indicator of the particular intervention that took place in each of the 125 municipalities. As long as whatever quality issues in the data apply to all municipalities, and we had no reason to believe this was not the case, we were on safe ground assuming the data fit our purposes.

HOMOGENEOUS TREATMENT EFFECTS ESTIMATES, A REPLICATION OF THE PROPENSITY SCORE AND MATCHING ESTIMATES

As mentioned before, as a first step to estimate the heterogeneity in treatment effects, we had replicated the matching procedure championed in Coneval’s evaluation: radius matching with a maximum distance of 0.05 for controls. Table 5 shows mean-comparison tests of several variables as balance diagnostics for comparing the distribution of baseline covariates between treatment groups.

Table 5
Mean-comparison tests between treatment groups
 (Radius caliper 0.05)

Variable	Unmatched		Mean		% bias	% reduct bias	t-test	
	Matched		Treated	Control			t	p > t
HDI 2005	U		0.62	0.76	-280.60		-26.18	0.00
	M		0.63	0.64	-21.10	92.50	-2.06	0.04
% of population under food poverty	U		68.45	31.00	277.50		23.45	0.00
	M		66.28	67.30	-7.60	97.30	-0.75	0.46
% of population under asset poverty	U		89.47	61.61	216.70		17.21	0.00
	M		88.54	89.46	-7.20	96.70	-1.14	0.26
Infant mortality rate 2005	U		36.87	22.39	200.80		20.48	0.00
	M		35.42	33.60	25.20	87.40	1.79	0.08
Female infant mortality rate 2005	U		32.65	19.83	200.80		20.48	0.00
	M		31.36	29.75	25.20	87.40	1.79	0.08
Scholar attendance rate 2005(6-24 years)	U		61.77	64.61	-47.80		-5.38	0.00
	M		63.11	63.77	-11.10	76.70	-0.69	0.49
Scholar female attendance rate 2005 (6-24 years)	U		1 228.60	3 752.90	-34.40		-2.68	0.01
	M		1 277.60	1 064.40	2.90	91.60	0.97	0.33
% of illiterate population of 15 years or more	U		41.98	15.37	271.60		30.27	0.00
	M		38.50	37.06	14.80	94.60	1.14	0.26
Average education level	U		3.59	6.18	-224.00		-19.17	0.00
	M		3.84	3.94	-8.80	96.10	-0.95	0.34
% of population over 5 years that speak indigenous language	U		7 443.70	2 159.80	69.70		9.14	0.00
	M		7 420.00	5 728.10	22.30	68.00	1.17	0.24
% of people living in houses with neither drainage system nor private toilet	U		18.71	9.85	47.80		7.56	0.00
	M		16.53	14.58	10.60	77.90	0.64	0.53

Table 5, continuation...

Variable	Unmatched	Mean		% bias	% reduct bias	t-test	
	Matched	Treated	Control			t	p > t
% of people living in houses without electricity	U	22.37	6.97	127.80		21.91	0.00
	M	19.85	14.44	44.90	64.80	2.50	0.01
% of people living in houses without tap water	U	52.54	19.22	149.10		18.12	0.00
	M	49.27	57.33	-36.10	75.80	-1.95	0.05
Logarithm of total population	U	8.82	9.35	-40.60		-3.69	0.00
	M	8.78	8.64	11.00	73.00	0.82	0.42
Logarithm of the total female population	U	8.16	8.68	-40.50		-3.68	0.00
	M	8.12	7.97	11.40	72.00	0.85	0.40
% of population from 0 to 4 years old	U	13.58	10.11	170.00		20.60	0.00
	M	12.99	13.39	-19.50	88.50	-1.19	0.24
% of population from 0 to 14 years old	U	42.51	33.13	206.70		20.68	0.00
	M	41.62	41.99	-8.00	96.10	-0.55	0.58
% of population of 60 years old or more	U	7.87	11.15	-82.90		-7.58	0.00
	M	8.59	8.20	10.10	87.90	0.80	0.43

Source: Own elaboration.

It's important to note that all variables exhibited major reductions in the standardized bias. Also in none of them was possible to reject the hypothesis of the equality of the means between treated and control after matching with a p -value smaller than 0.01, and only in two cases with p -value smaller than 0.05. Moreover, there was added value in using exactly the same matching as the one in Coneval's evaluation, for it allowed us to reexamine Coneval's conclusions in its own terms, providing evidence of heterogeneous treatment effects of the E100×100 program, something that remained a conjecture in the original evaluation.

Remember that these matching results were based on the ITM which means that the original assignment of the municipalities to E100×100 was used to carry out the analysis instead of the effective reception of the treatment, that in this case would've correspond to the actual degree of coordination achieved by the different institutions part of the E100×100.

Also of importance is the fact that the treated municipalities that were unmatched using the radius matching algorithm of 0.05 are the ones which have the worst conditions among all the treated ones. This explains why it was not

possible to find a good match for them since they were the most disadvantaged municipalities among the rest that were part of the E100×100 program. Hence, let's keep in mind that what follows, as Coneval's conclusions, is based only on the 73 treated municipalities for which a good match was found.

At this point we used the level of investment per capita received through the program's specific actions as a sign of the mixed nature of E100×100's intervention to uncover any underlying systematic heterogeneity in the effects of the E100×100 program.

As the annual investment data show, municipalities in the E100×100 program differed greatly in the particular "treatment" they received as the amount of investment per capita varies across municipalities. Provided the amount of investment per capita reflects differences in the intervention orchestrated by E100×100, it was to be expected that municipalities whose treatment included the required financial resources would observe the expected effects.

HETEROGENEOUS TREATMENT EFFECTS ESTIMATES

The hypothesis of our paper is that, provided the different levels of investment per capita by specific action among the 125 treated municipalities reflects actual differences in E100×100's intervention, municipalities differ not only in their treatment status but also in the particular treatment they received. Thus, a systematic heterogeneity in the treatment effects estimated by Coneval was expected to be shaped depending on whether the intervention was accompanied by the resources dearly needed in the treated municipalities.

In order to estimate the heterogeneous treatment effects as a function of the investment per capita, for every outcome variable, the observed difference in a pair between a treated municipality and the group mean of the multiple matched controls municipalities was estimated. That is, following Xie, Brand, and Jann (2012), the data was transformed so that the differences in pairs between matched treated and control municipalities constituted the "observed" data subject to further modeling. In other words, the "raw" data for the next step were observational differences of matched comparisons for each outcome variable in this third step.

These differences were used to run a parametric model (Ordinary Least Squares) to assess the statistical significance of a linear trend of the treatment effect heterogeneity as a function of the associated actions' investment per

capita level.⁸ Of course, more flexible modeling devices could be used to fit the data such as locally weighted regression. We also followed this approach, which we don't show here for reasons of space, and the results suggest the linear trend as a good approximation.

Models with a significant linear trend were analyzed to answer whether or not, when E100×100's intervention included a minimum of financial resources, the impact estimated for those particular municipalities were the desired ones.

Whenever a statistical significance linear trend was found in the analysis of every action associated to the outcome variable subject of study, the differences between the treated municipalities and the group mean of the multiple matched controls were plotted along the level of investment per capita—in the *x*-axis. This allowed us to analyze the relation between the observed differences of matched comparisons—for each of the 29 outcome variables—and the investment per capita by specific action in the treated municipalities (the heterogeneous treatment effects).

Additionally as Xie, Brand, and Jann (2012) suggest, the treatment effects may vary systematically by the propensity score for treatment due to heterogeneity population composition. Hence, the observed matched differences were also plotted against the propensity score in order to study the heterogeneity in treatment effects as a function of the probability of receiving treatment. The propensity score itself was an index of social exclusion in this case. This was done in order to find evidence on which treated municipalities benefited more from E100×100 and therefore provided evidence that could be used by policy makers in order to assign resources more effectively.

RESULTS

For reasons of space, and to avoid flooding the reader with numbers, this section presents only part of the results of the regression analysis of the matched differences against investment per capita disaggregated by specific actions. Only those actions that were expected to have a direct impact on the outcome variable in question were reported in this section. We've also omitted the results

8 Another possibility is to estimate the statistical significance of an interaction term in a weighted regression of the outcome variable on the treatment indicator and the investment per capita using as weights the frequency with which the observation is used as a match. However, here we follow Xie and colleagues' approach.

for the outcome variables for which no statistically significant —robust errors were taken into account whenever they were bigger than the usual errors— relations were found.⁹

Table 6 shows the main results of the analysis of treatment heterogeneity for the percentage of illiterate population of 15 years and older in the municipality.

Table 6
Heterogeneity analysis for illiteracy

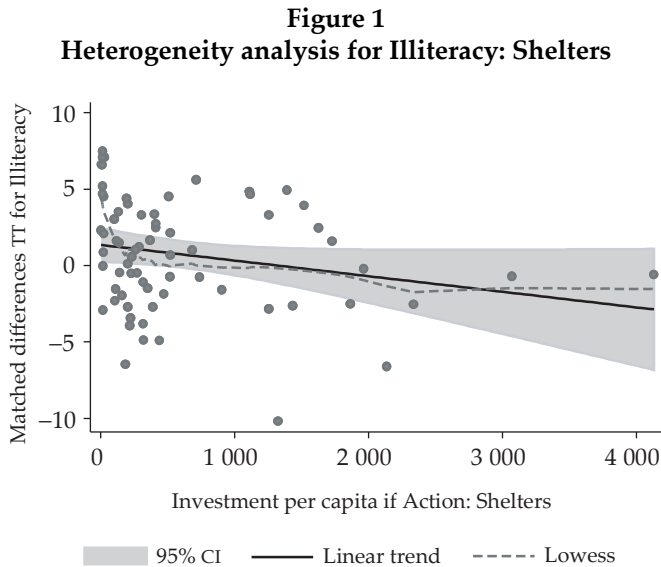
	<i>Matched differences ATT for illiteracy</i>			
	Levels 10	First differences	Double differences	Triple differences
Shelters	-0.00149* (-0.0007)	-0.00140** (0.000629)	-0.00236 (0.00144)	-0.00201** (0.000963)
Observations	69	53	48	48
R-squared	0.022	0.088	0.055	0.086
Propensity Score	8.553** (3.259)	0.568 (1.553)	7.586** (3.692)	-0.186 (2.654)
Observations	73	73	65	65
R-squared	0.088	0.002	0.063	0.000

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. First difference: illiteracy 2010- illiteracy 2005; Double differences: illiteracy 2010-(illiteracy 2005-illiteracy 2000); Triple differences: illiteracy 2010-illiteracy 2005-(illiteracy 2005-2000).
Source: Own elaboration.

From these results it can be seen that the coefficient of investment per capita for the action of Shelters¹⁰ is statistically significant and negative for three of the estimated differences: the simple difference taken in 2010 on the matched municipalities; first differences, that is the difference in difference 2010-2005 and the difference in difference that discounts the difference in the observed trend 2005-2000. All of this suggests that there was a positive impact on the reduction of the expected illiteracy as the level of investment per capita for Shelters actions increased.

9 The bulk of the results are available from the authors on request.
10 This strategic includes specific actions such as educative packages delivered to Shelters; infrastructure actions in the Shelters; Scholars in Shelters and actions performed in assistance of the indigenous population.

This is illustrated in Figure 1 where the difference between a treated municipality and the group mean of the multiple matched controls is plotted against the investment per capita in Shelters. The figure also depicts the *lowess* or fully nonparametric estimation for the same variables and it seems that for this specific case linearity is a reasonable functional form.



Source: Own elaboration.

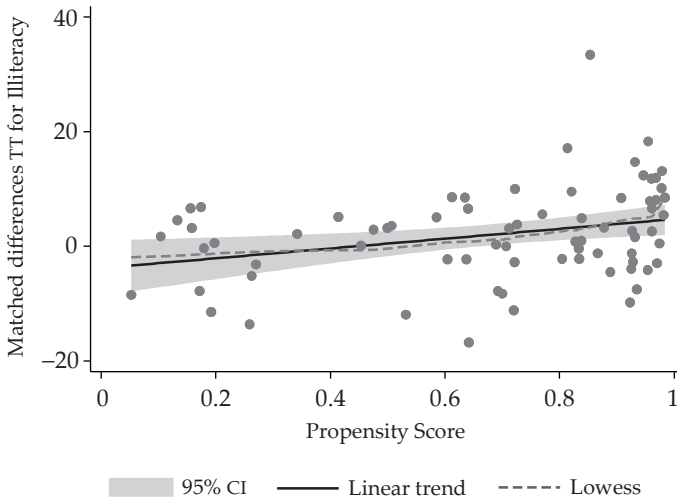
These results, we argue, were informative enough to assert that in those treated municipalities where the intervention was accompanied by investment in Shelters, the impact on the reduction of illiteracy was larger than in those municipalities in which the intervention didn't secure a minimum of financial resources. It is in this heterogeneity that the estimated impact in Coneval's evaluation got lost.

Table 6 also presents the results of the analysis where the heterogeneous treatment effects of the program were estimated as a function of the ps to treatment. The coefficient of the ps was statistically significant and positive for the simple difference in levels and the double difference 2010-(2005-2000).

Thus, as it can be seen in figure 2, the expected impact of E100×100 on the reduction of illiteracy was largest for the municipalities that had a lower ps—and therefore less likely to be assigned into the program—and attenuates as

poorer municipalities—the ones with greater propensity scores and therefore more vulnerable—were considered.

Figure 2
Heterogeneity analysis for Illiteracy: ps



Source: Own elaboration.

This implies that the municipalities that have a greater PS—the ones who are most likely to be part of the program—are the ones who are benefiting less from it. Together, these results provided compelling quantitative evidence that the E100×100 program has proven most helpful in less vulnerable territories, leaving behind those in greater need.

Table 7 shows the results for the Rate of Scholar Attendance from 6 to 24 years old. The results indicated a statistically significant trend on the improvement of the rate of scholar attendance as the level of investment per capita increased for the strategic actions Educative Packages and Rural Roads for the simple difference in 2010, other actions in Education¹¹ for the difference in difference, and Cement Flooring¹² for the simple difference, double and triple

¹¹ This includes actions such as Assistance for indigenous education; equipped libraries; Elementary school services among many others

¹² This consists on the provision of cement floors in the households that do not have it.

difference. Again, this table shows that the improvement in the rate of scholar attendance was greater for those municipalities with greater investment. In other words, what these figures show is that there was a progressively increasing positive effect on the rate of scholar attendance as the level of investment per capita in the actions mentioned above increased.

Table 7
Matched differences ATT for scholar attendance

<i>Actions</i>	<i>Levels 2010</i>	<i>First differences</i>	<i>Double differences</i>	<i>Triple differences</i>
Educative Packages	0.369*** (0.130)	-0.107 (0.0912)	0.203 (0.181)	-0.156 (0.173)
Observations	73	56	51	51
R-squared	0.141	0.025	0.025	0.016
Others Education	0.00226 (0.00653)	0.00996* (0.00556)	-0.00999 (0.0101)	-0.00362 (0.00968)
Observations	73	56	51	51
R-squared	0.001	0.056	0.020	0.003
Rural Roads	0.0003** (0.0001)	-0.000219 (0.000146)	0.000134 (0.000284)	-0.000185 (0.000276)
Observations	68	52	47	47
R-squared	0.035	0.043	0.005	0.010
Cement Flooring	0.0031* (0.00150)	0.00151 (0.00105)	0.00493** (0.00242)	0.00464* (0.00232)
Observations	70	54	49	49
R-squared	0.080	0.038	0.081	0.079
Propensity Score	-5.179** (2.029)	-2.900* (1.614)	-0.906 (2.735)	0.917 (2.788)
Observations	73	73	65	65
R-squared	0.049	0.044	0.002	0.002

Note: The ols regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. First difference: (2010-2005); Double differences: 2010-(2005-2000); Triple differences: (2010- 2005)-(2005-2000).
Source: Own elaboration.

The results for the heterogeneous treatment effects of the program as a function of the ps to treatment suggested that the improvement in the rate of scholar attendance was larger for those municipalities who were less likely to be part of it.

Table 3 presents the results for the Rate of Scholar Absence from 6 to 14 years old. The results were statistically significant and negative for the investment per capita in Educative Packages for the simple difference in 2010 and Cement Flooring for the simple difference of 2010, double and triple differences. This indicated that as the level of investment per capita in each of these actions increased, the estimated impact on the reduction on the rate of scholar absence increased as well. Thus, the expected effect of E100×100 on the Rate of Scholar Absence was larger for those municipalities on which the intervention included financial resources to the municipalities through these specific actions. It is worth mentioning that the specific action Cement Flooring of the Housing Aspect was believed to have an impact on the reduction of the scholar absence as it provided households with better hygienic conditions and therefore reduced the probability of getting sick and skip school.

Table 8
Matched differences ATT
for Rate of Scholar Absence 6 to 14 years old

<i>Actions</i>	<i>Levels 2010</i>	<i>First differences</i>	<i>Double differences</i>	<i>Triple differences</i>
Educative Packages	-0.190** (0.0796)	0.0526 (0.0534)	-0.133 (0.142)	0.105 (0.105)
Observations	73	56	51	51
R-squared	0.074	0.018	0.018	0.020
Cement Flooring	-0.00120 (0.000893)	-0.000839 (0.000674)	-0.00405** (0.00190)	-0.00254* (0.00142)
Observations	70	54	49	49
R-squared	0.022	0.029	0.088	0.064
Propensity Score	6.591*** (1.799)	-0.570 (1.035)	3.266 (2.577)	-2.865 (2.114)
Observations	73	73	65	65
R-squared	0.159	0.004	0.025	0.028

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. First difference: (2010-2005); Double differences: 2010-(2005-2000); Triple differences: (2010- 2005)-(2005-2000).

Source: Own elaboration.

The results of the heterogeneity in the treatment effects of E100×100 as a function of the PS show that the coefficient of the PS was statistically significant and inversely correlated to the estimated effect on the Rate of Scholar Absence.

Again, this relationship showed once more a progressively increasing effect, this time on the reduction of the rate of scholar absence, as municipalities with lower PS were included.

Table 9 illustrates the results for the percentage of households without washing machines. These results indicated that there was a progressively greater effect as the investment per capita in actions of Training, Savings and Financing—in simple and double differences—and Support Programs for Economic Activities—in simple, first and double differences—increased.

Table 9
Matched differences ATT for percentage
of households with no laundry

<i>Actions</i>	<i>Levels 2010</i>	<i>First differences</i>	<i>Double differences</i>	<i>Triple differences</i>
Training, Saving and Financing	-0.012*** (0.00390)	-0.00539 (0.00345)	-0.0099** (0.00457)	-0.000664 (0.00349)
Observations	73	56	51	51
R-squared	0.126	0.043	0.088	0.001
Support Programs	-0.00293*** (0.00105)	-0.00207** (0.000823)	-0.003*** (0.00110)	-0.000729 (0.000877)
Observations	73	56	51	51
R-squared	0.106	0.105	0.181	0.014

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. First difference: (2010-2005); Double differences: 2010-(2005-2000); Triple differences: (2010- 2005)-(2005-2000).

Source: Own elaboration.

Similar results were also found for the percentage of households without refrigerators for the first, double and triple differences in the case of Support Programs and for the double differences for Training, Savings and Financing.

The last tables show the estimated results on Coneval's Multidimensional Measurement of Poverty, variables only available for the year 2010, this being the reason why no estimations are presented under the difference in difference approach.

Table 10 shows the results for the percentage of people in poverty due to quality and spaces of the dwelling. This table shows statistically significant and negative coefficients for the investment per capita in the actions of Stoves¹³

¹³ Which entail activities such as the installation and use of stoves.

Temporal Employment and Basic Unit Housing. That is, a progressively increasing impact was observed as the investment per capita in these actions increased. Again what these figures show is that municipalities with greater levels of investment per capita in these actions experienced a larger reduction in the percentage of people in poverty due to quality and spaces of the dwelling outcome indicator.

Table 10
Matched differences ATT for deprived
in the quality and spaces of housing

<i>Actions</i>	<i>Levels 2010</i>	<i>Actions</i>	<i>Levels 2010</i>
Temporal Employment	-0.0198* (0.0115)	Stoves	-0.0339*** (0.0117)
Observations	60	Observations	72
R-squared	0.025	R-squared	0.070
Propensity Score	-17.78*** (6.229)	Basic Unit Housing	-0.00359** (0.00162)
Observations	73	Observations	45
R-squared	0.103	R-squared	0.049

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
 Source: Own elaboration.

The coefficient estimated for the trend in the PS was statistically significant indicating that the municipalities that were more likely to participate in the program were the ones which had the largest improvement in this indicator. A result that contrasted with that of the educational aspect of the program, suggesting the different contexts in which these actions seem to work for the benefit of the targeted population.

Table 11 presents the results for the percentage of the population not multidimensional poor and not vulnerable. Two specific actions showed a statistically significant trend: Cement Flooring and Schools. As shown in this table, there was a gradually increasing positive effect on this indicator as the level of investment per capita in the actions increased.

For the case of the PS, the results suggest that the impacts of the program were largest for those municipalities that had PS close to unity. Meaning that there was a progressively increasing effect of the program as the poorest municipalities were included.

Table 11
Matched ATT differences
for not poor neither vulnerable population

<i>Actions</i>	<i>Levels 2010</i>	<i>Actions</i>	<i>Levels 2010</i>
Propensity Score	0.512*** (0.143)	Cement Flooring	0.00014* (8.3e-05)
Observations	73	Observations	70
R-squared	0.154	R-squared	0.058
Schools	0.0002** (0.0001)		
Observations	73		
R-squared	0.097		

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: Own elaboration.

Table 12 shows the results for the outcome variable population with at least three social deprivation indicators. The coefficients for the Cement Flooring and Others in Education actions were statistically significant and negative, which means that the reduction on these outcome variables increased along with investment per capita in these actions.

Table 12
Matched ATT differences for population
with at least three social deprivation

<i>Actions</i>	<i>Levels 2010</i>	<i>Actions</i>	<i>Levels 2010</i>
Others in Education	-0.0271** (0.0113)	Cement Flooring	-0.00447* (0.00227)
Observations	73	Observations	70
R-squared	0.079	R-squared	0.062

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: Own elaboration.

Lastly, in Tables 13 and 14 we see the results for the percentage of population with income below the minimum welfare line indicator and extreme poverty. In Table 13, the investment in three specific actions showed a statistically significant trend: Support Programs for Economic Activities¹⁴, Literacy, and Shelters.

¹⁴ Which include specific actions such as Support for farmers and artisans, and mothers assisted by the daily childcare center.

In Table 14, the coefficients for the investment per capita in Energetic Support, Food Assistance¹⁵ and Temporal Employment were statistically significant and negative. Indicating again that the municipalities that received the largest amounts of investment on these actions were the ones who exhibited greater reductions of this indicator.

Table 13
Matched ATT differences for population
with income below the minimum welfare line

<i>Actions</i>	<i>Levels 2010</i>	<i>Actions</i>	<i>Levels 2010</i>
Shelters	-0.00138* (0.000771)	Support Programs	-0.00181* (0.000927)
Observations	69	Observations	73
R-squared	0.031	R-squared	0.042
Literacy	-0.0112** (0.00464)		
Observations	73		
R-squared	0.092		

Note: The OLS regressions were run with robust standard errors and standard errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
 Source: Own elaboration.

Table 14
Matched ATT differences for extreme poverty

<i>Actions</i>	<i>Levels 2010</i>	<i>Actions</i>	<i>Levels 2010</i>
Temporal Employment	-0.0125* (0.00664)	Energetic Support	-0.0111* (0.00662)
Observations	60	Observations	73
R-squared	0.037	R-squared	0.031
Food Assistance	-0.0022* (0.00123)		
Observations	73		
R-squared	0.033		

Note: The OLS regressions were run with robust standard errors and standard Errors being the larger errors the ones who are reported in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
 Source: Own elaboration.

15 Provision of food supplements; Fortified milk delivered to households and families with grocery support.

CONCLUSIONS

In the pursuit of effectiveness of development policies, the crucial need to coordinate different levels of governments and institutions is not a new topic and its importance is continuously advocated (see Bouckaert, Peters, and Verhoest, 2010; Peters, 2015).

Since the seventies Mexico has ventured in the design of policy strategies oriented at solving the shortcomings left by poor management of interdependencies in development policies. Indeed, the efficiency costs of the lack of coordination and collaboration between players at different levels of government, ministries and program administrators has long been recognized.

Every new administration endorses the need to develop a comprehensive plan to determine and respond to the needs for coordinating processes and procedures among adjacent institutions and between levels of government. However, managing this kind of interrelations has proven an elusive undertaking. Admittedly, this is not an easy task. It is a complex web of interdependencies, simultaneously vertical, across different levels of government, and horizontal, among ministries and programs at the same level of government, and networked (Koliba, Meek, and Zia, 2010). Surprisingly, there is only a handful of studies in Mexico that investigate the tools being used to help bridge the gaps that exist between levels of government as they design and implement public policies (see Collins *et al.*, 2004; Ordaz Ocampo, 2012; Camacho García and Flamand, 2008; Cejudo Ramírez, 2012; Kroeger and Luna, 1992; Yashchine, Ochoa, and Hernández, 2014).

From our part, at the results level, the evidence presented in this paper suggests a positive impact of E100×100 whenever managed to get accompanied by a minimum of investment. Hence, it's possible to give an affirmative answer to the hypothesis that whenever E100×100's intervention included the needed investment per action to the treated municipalities, the expected effects were found. This shows that even though, in the intention-to-treat evaluation carried out by Coneval, it was not possible to find many statistical significant differences, this did not imply that the intervention was not effective. It simply means that the E100×100 couldn't intervene equally across municipalities. Furthermore, the evidence presented suggests the expected results were observed for those outcome variables whose associated actions successfully secure a minimum of investment.

These findings bring new elements of analysis for the intention-to-treat evaluation done by Coneval indicating that it was less likely that their evaluation did not find the expected effects because the control municipalities ended up being treated. On the contrary, it has been shown that it is more likely that the municipalities in the treatment group actually received different treatments, some more effective than others. In other words, it has been suggested that E100×100's intervention seems to have worked best whenever it managed to get accompanied by a certain level of investment. However, in more cases than not, the E100×100 umbrella program couldn't go that last mile.

For the actions on which the intervention did not secure a minimum investment, the results were not the expected ones. Again, the important thing to notice is that the evidence provided suggests that the intention-to-treat evaluation did not find many statistical differences because the municipalities in the treatment group actually ended up not being treated differently than the control group in the sense that the intervention did not always include financial resources. Moreover, the evidence presented here poses severe doubts on the hypotheses advanced in the previous evaluation that no significant results were found due to the possibility that some of the municipalities in the control group were also intervened.

What is behind these differences in the treatment received by the municipalities is hard to know without field data. However, it has long been recognized that the interinstitutional coordination is a complex matter and that policy designers face several hurdles to establish common goals and have an efficient communication among the actors involved. Moreover, the task of setting strategic priorities is inherently political and this could also pose several limits to the effectiveness of coordinated strategies (Evans, 1995). It has also been known for a long time now that a key element in the efforts to develop an efficient and sustainable coordination of public service delivery is the existence of a series of formalised measures, mechanisms and steps that ensure the commitment in the activities that need to be carried out by the actors involved in coordination process (Grindle, 2010; North and Shariq, 2004; Ordaz Ocampo, 2012). An element missing in the original design of E100×100.

Hardly anyone would contest the importance of competent and committed individuals with strong values, good judgment and sound knowledge and understanding of the mutual dependence of policy responsibilities behind it

all. However, without the management basis for the entire interinstitutional coordination plan, a document that establishes committees and policies placing responsibility and authority where they are needed to accomplish tasks, there is little hope even for otherwise well-intentioned officials to overcome the natural resistance of competition and turf protection (Yaschine, Ochoa, and Hernández, 2014).

Thus, the Mexican government needs to start taking actions in order to develop and implement the formal mechanisms that represent a challenge for the interinstitutional coordination. The mechanisms should focus on the clear recognition of the existing government conditions and capacities to set and implement feasible priorities in a strategic manner. There is a need for clarity in roles, responsibilities and objectives. The need to focus on the specific requirements of the communities in which services are delivered is apparent. The search for diverse methods to ensure accountability and the creation of regulatory frameworks or agreements for collaboration to take place are a must. A strong national leadership and management is required to strengthen both local government and civil society. The encouragement of participation between communities and institutions to provide local solutions to local issues, along with regular meetings and effective communication among the representatives from many different ministries and agencies across government and the investment of time and resources into community consultations and consensus building are needed (Keast, 2011; Stewart, Lohar, and Higgins, 2011; Grindle, 2010; Gray, 2002; Leigh, 2008; Grindle, 2004; Ordaz Ocampo, 2012; Sabatier, 1986).

As long as the Mexican government continues assigning economic resources to programs like the E100×100 without creating the necessary conditions mentioned above, any further attempts such as the recent *Cruzada Nacional Contra el Hambre* [National Crusade Against Hunger] launched by the current administration will continue most likely to not show the desired results. Nowadays, Coneval has commissioned a specific assessment of the interinstitutional coordination processes and procedures behind *Cruzada Nacional Contra el Hambre*, a program that certainly left things to be desired from the start (Yaschine, Ochoa, and Hernández, 2014). The results of which, expected by 2016, will likely point in the same direction.

REFERENCES

- Bernal, R., and Peña, X., 2011. *Guía práctica para la evaluación de impacto*. Bogota, Colombia: Universidad de los Andes, Facultad de Economía, Centro de Estudios sobre Desarrollo Económico.
- Bouckaert, G., Peters, B.G., and Verhoest, K., 2010. *The coordination of public sector organizations: shifting patterns of public management*. Palgrave Macmillan.
- Brand, J.E., and Thomas, J.S., 2013. Causal effect heterogeneity. In: S.L. Morgan (ed.). *Handbook of Causal Analysis for Social Research* (pp. 189-213). New York: Springer Netherlands.
- Brown, D., and Kalegoankar, A., 2000. *Intersectoral Cooperation: Lesson for Practice*. Boston: Institute for Development and Research.
- Brown, D., and Tandon, R., 1992. *Public-NGO Financing Institutions in India: An Exploratory Study*. Washington, DC: The World Bank.
- Camacho García, M., and Flamand, L., 2008. Políticas intergubernamentales para controlar la contaminación del aire en ciudades mexicanas: Una evaluación. *Gestión y política pública*, 17(2), pp. 261-313.
- Cejudo Ramírez, G.M., 2012. Las políticas de gestión pública en México: valorando 12 años de reformas en democracia. *XVII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública*, Cartagena, Colombia, October 30-November 2. Available at: <[http://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/01B97A3EA1AD9E4705257C38006E95B0/\\$FILE/cejudram.pdf](http://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/01B97A3EA1AD9E4705257C38006E95B0/$FILE/cejudram.pdf)>.
- Collins, K., Ganster, P., Mason, C., Sánchez López, E., and Quintero-Núñez, M., 2004. *Imperial-Mexicali Valleys: Development and environment of the US-Mexican Border Region*. San Diego, CA: San Diego State University Press.
- CONEVAL (Consejo Nacional de Evaluación de la Política de Desarrollo Social), 2005a. *Índice de rezago social*. Mexico, D.F.: CONEVAL.
- CONEVAL, 2013. *Evaluación de impacto de la Estrategia 100x100*. [online] Mexico, D.F.: CONEVAL. Available at: <http://web.coneval.gob.mx/Informes/Evaluacion/Impacto/Evaluacion_de_impacto_de_la_Estrategia_100x100.pdf> [Accessed 20 March 2014].
- CONAPO (Consejo Nacional de Población), 2005. *Índice de marginación 2005*. Mexico, D.F.: CONAPO.
- Evans, P., 1995. *Embedded Autonomy: States and Industrial Transformation*. Princeton, NJ: Princeton University Press.

- Gertler, P.J., Martínez, S., Premand, P., Rawlings, L.B., and Vermeersch, C.M., 2011. *Impact Evaluation in Practice*. Washington, DC: The World Bank. Available at: <http://siteresources.worldbank.org/INT/HDOFFICE/Resources/IEP_SPANISH_FINAL_110628.pdf>.
- Gray, A., 2002. *Integrated service delivery and regional co-ordination: a literature review*. Wellington, NZ: Gray Matter Research Ltd.
- Gray, B., 1989. *Collaborating: Finding Common Ground for Multiparty Problems*. San Francisco, CA: Jossey-Bass Publishers.
- Grindle, M., 2004. Good Enough Governance: Poverty Reduction and Reform in Developing Countries. *Governance: An International Journal of Public Policy and Administration*, 17(4), pp. 525-548.
- Grindle, M., 2010. Social Policy in Development: Coherence and Cooperation in the Real World [DESA Working Paper no. 98 ST/ESA/2010/DWP/98]. *United Nations Department for Economic and Social Affairs*, New York.
- Guo, S., and Fraser, M.W., 2014. *Propensity Score Analysis: Statistical Methods and Applications* [Vol. 11]. Thousand Oaks, CA: Sage Publications, Inc.
- Herrera, J., 2007. La evaluación en México. In: M. Sánchez Silva, C.M. Valenzuela Espinoza, and J.L. Sánchez Casténum (comps.). *Seminario Internacional. Gobierno y Políticas Públicas (Primeras Propuestas) Gobierno del Estado de Sinaloa, México*. Mexico: Instituto Politécnico Nacional, Centro de Investigaciones Económicas, Administrativas y Sociales (CIECAS).
- INEGI (Instituto Nacional de Estadística y Geografía), 2000. *XII Censo General de Población y Vivienda 2000*. Mexico: INEGI.
- INEGI, 2005. *II Conteo de Población y Vivienda 2005*. Mexico: INEGI.
- INEGI, 2010. *Censo de Población y Vivienda 2010*. Mexico: INEGI.
- ITESM (Instituto Tecnológico y de Estudios Superiores de Monterrey), 2007. *Evaluación de la estrategia de microrregiones: reporte de diseño*. [online] Mexico: ITESM. Available at: <https://inapam.gob.mx/work/models/SEDESOL/Resource/3156/1/images/U_ReportedeDiseno.pdf> [Accessed 20 March 2014].
- Keast, R., 2011. Joined-up governance in Australia: How the past can inform the future. *International Journal of Public Administration*, 34(4), pp. 221-231.
- Koliba, C., Meek, J.W., and Zia, A., 2010. *Governance Networks in Public Administration and Public Policy*. Boca Raton, FL: CRC Press Taylor & Francis Group.
- Kroeger, A., and Luna, R. (comps), 1992. *Atención primaria de salud. Principios y métodos*. Washington, DC: Organización Panamericana de la Salud.
- Leigh, J., 2008. *Improved integration and coordination of services: evaluation of the Stronger Families and Communities Strategy 2000-2004*. [online] Canberra: Department of

- Families, Housing, Community Services and Indigenous Affairs. Available at: <<http://mams.rmit.edu.au/gca34vrtfd76.pdf>> [Accessed 15 June 2014].
- Lustig, N., 2001. Life is not easy: Mexico's quest for stability and growth. *The Journal of Economic Perspectives*, 15(1), pp. 85-106.
- Manski, C., 2007. *Identification for Prediction and Decision*. Cambridge, MA: Harvard University Press.
- North, D., and Shariq, S., 2004. Learning, institutions, and economic performance. *Perspectives on Politics*, 2(1), pp. 75-84.
- Ordaz Ocampo, J., 2012. *La Estrategia 100×100. Hacia la elaboración de un modelo de implementación de políticas públicas de atención integral*. Tesis de Maestría. Facultad Latinoamericana de Ciencias Sociales (Flacso).
- Outhwaite, W., and Turner, S. (eds.), 2008. *The SAGE Handbook of Social Science Methodology*. London: SAGE Publications Ltd.
- Pan, W., and Bai, H. (eds.), 2015. *Propensity Score Analysis: Fundamentals and Development*. New York: The Guilford Press.
- Peters, B.G., 2015. *Pursuing Horizontal Management. The Politics of Public Sector Coordination*. Lawrence, KS: University Press of Kansas.
- PNUD (Programa de las Naciones Unidas para el Desarrollo), 2000. *Índice de Desarrollo Humano 2000*. Mexico: PNUD.
- PNUD, 2005. *Índice de Desarrollo Humano 2005*. Mexico: PNUD.
- Sabatier, P., 1986. Top-down and bottom-up approaches to implementation research: a critical analysis and suggested synthesis. *Journal of public policy*, 6(01), pp. 21-48.
- Stewart, J., Lohar, S., and Higgins, D., 2011. *Effective Practices for Service Delivery Coordination in Indigenous Communities* (Closing the Gap Clearinghouse Resource Sheet no. 8). Canberra: Australian Institute of Health and Welfare and Australian Institute of Family Studies.
- Xie, Y., 2013. Population heterogeneity and causal inference. *Proceedings of the National Academy of Sciences*, 110(16), pp. 6262-6268.
- Xie, Y., Brand, J., and Jann, B., 2012. Estimating heterogeneous treatment effects with observational data. *Sociological Methodology*, 42(1), pp. 314-347.
- Yaschine, I., Ochoa, S., and Hernández, C., 2014. Cruzada Nacional Contra el Hambre: Análisis de su diseño desde un enfoque de seguridad alimentaria [Documento de trabajo no. 10]. *Programa Universitario de Estudios del Desarrollo, Universidad Nacional Autónoma de México* (PUED, UNAM), Mexico, D.F. Available at: <<http://www.pued.unam.mx/archivos/opinion/010.pdf>>.