

# Manipulation of Social Program Eligibility

by Adriana Camacho and Emily Conover

## Web Appendix

### A. Manipulation and Costs of Cheating

In this Appendix, we estimate the relationship between the number of surveys conducted and amount of cheating observed with the costs of cheating. The available data that proxies for the cost of cheating do not vary over time. We use number of community organizations and number of the main newspaper in circulation as measures for the costs of manipulation in a given municipality.

DATA. — Previous studies have found that institutions are stronger when there is more community oversight or when citizens are better informed.<sup>1</sup> One measure for institution quality that we use is the number of community organizations in each municipality in 1998. Rosas and Mendoza (2005) describe community organizations as neighborhood level government accountability and conflict resolution entities sometimes involved in local infrastructure projects. An article by Sandra V. Chávez (2006)<sup>2</sup> explains the key role that community organizations play in improving and monitoring the transparency of public resources. These data come from a non-profit civil foundation, the Social Foundation (*Fundación Social*).

We also use newspaper circulation data, with the idea that it is harder to cheat in municipalities where the citizens are better informed about public affairs.<sup>3</sup> Newspaper circulation corresponds to certified daily average circulation data by municipality for 2004 from Colombia's main national newspaper, *El Tiempo*.

Other cross section data that we use include: an alternative measure for poverty in a municipality which is the proportion of people with Unsatisfied Basic Needs (NBI in Spanish) constructed using information from the 1993 and 2005 Population Census; the distance from the municipality to the largest city in the state measured in kilometers; and the size of the municipality in square kilometers. These data come from DANE. Summary statistics are provided in the main text in Table 2.

<sup>1</sup>Ariel Fiszbein (1995); Timothy Besley and Robin Burgess (2002); Andrés Rosas and Juan Mendoza (2005)

<sup>2</sup>"Transparencia, Antídoto contra la Corrupción", July 2006, available at: <http://www.comunikandonos.com/sitio/reputacion-coporativa/78-transparencia-antoto-contra-la-corrupci.html>

<sup>3</sup>See Sam Schulhofer-Wohl and Miguel Garrido (2009) for a recent article on how a decrease in newspaper circulation affected the number of candidates that ran for office, the probability of the incumbent winning an election and voter turnout in Cincinnati.

EMPIRICAL FINDINGS. — The equation we use to determine whether the proportion of surveys is smaller when the costs of cheating are higher has the following form:

$$(A1) \quad \begin{aligned} \textit{proportion\_surveys}_{jt} &= \alpha + \beta_1 \textit{costs}_j + \beta_2 \textit{lnpop}_{jt} + \beta_3 \textit{demography}_{jt} \\ &+ \beta_4 \textit{geography}_j + \eta_t + \epsilon_{jt} \end{aligned}$$

Where the dependent variable *proportion\_surveys*, is the fraction of annual surveys conducted in the 6 months prior to the election. *costs* is either the number of community organizations or the daily average newspaper circulation from May to October in each municipality.<sup>4</sup>

The equation we use to determine whether the size of the discontinuity is smaller when the costs of cheating are higher has the following form:

$$(A2) \quad \begin{aligned} \textit{discontinuity}_{jt} &= \alpha + \beta_1 \textit{costs}_j + \beta_2 \textit{lnpop}_{jt} + \beta_3 \textit{demography}_{jt} \\ &+ \beta_4 \textit{geography}_j + \eta_t + \epsilon_{jt} \end{aligned}$$

Where the dependent variable *discontinuity*, again proxies for the amount of cheating in a municipality. *costs* are defined as in equation A1.

A concern about running a cross section regression is that the mode of the score distribution is centered at a different point for each municipality depending on its wealth level. In the variable *demography* we control for differences in poverty rates across municipalities by including a measure of the proportion of people with unsatisfied basic needs calculated from the 1993 and 2005 population census. We also included in all regressions a measure of the size of the population *lnpop*, and the proportion of urban population in each municipality. To control for the possibility that more remote areas could have more cheating because of weaker presence of the state, we included the distance to the largest city in the *departamento* (state). Also in the *geography* variable we included the surface area of the municipality. We expect to see that municipalities with better monitoring institutions have less cheating. For these regressions we used the same years that we used in the main results: 1997, 2000 and 2003.

We report standardized results. We find that the coefficients have the expected signs, consistent with the idea that better monitoring is associated with a lower fraction of surveys in the 6 months before the election and less cheating in municipalities around election times. Columns (1)-(4) of Table A1 indicate that a standard deviation increase in the number of community organizations or newspaper circulation (s.d. = 325 and 3154 respectively) are associated with a lower fraction of surveys conducted before the elections of 0.03 of a standard deviation. Columns (1)-(4) of Table A2 use the fraction of surveys three points below and above the threshold, while columns (5)-(8) use the fraction of surveys 5 points below and above the threshold. The results indicate that a standard deviation increase in the number of community organizations (newspaper

<sup>4</sup>We use data from May to October because this would correspond to the period six months before the mayoral elections.

circulation) are associated with a lower percent of interviews three points below the threshold relative to three points above the threshold of 0.026 (0.035) of a standard deviation. The corresponding numbers for five points below the threshold are 0.05 and 0.06 for community organization and newspaper circulation respectively. Also, consistent with the model prediction we see that there is an inverse relation between the discontinuity at the threshold and the proportion of poor in a municipality in Table A2.

TABLE A1—NUMBER OF SURVEYS AND COSTS OF CHEATING (CROSS SECTION)

Dependent variable:	Proportion of surveys			
	(1)	(2)	(3)	(4)
Number of community organizations	-0.032*** [0.010]	-0.034*** [0.010]		
Newspaper circulation			-0.034*** [0.012]	-0.035*** [0.012]
Proportion of poor	0.030 [0.031]	0.049 [0.037]	0.022 [0.039]	0.046 [0.046]
Log Population	0.050 [0.035]	0.060* [0.036]	0.046 [0.045]	0.048 [0.047]
Year effects	Yes	Yes	Yes	Yes
Demography controls	Yes	Yes	Yes	Yes
Geography controls		Yes		Yes
Observations	1295	1295	853	853
R-squared	0.07	0.07	0.07	0.07

Note: Robust standard errors in brackets. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. All regressions include an intercept term and report standardized results. The dependent variable is the fraction of annual surveys conducted in the 6 months prior to the election during election years. Demography controls include proportion urban population and proportion poor population using a measure for unsatisfied basic needs from the 1993 and 2005 Population Census. Geography controls include distance to the largest city (capital) in the *departamento* (state) and municipality surface area.

TABLE A2—DISCONTINUITY AT THE THRESHOLD AND COSTS OF CHEATING (CROSS SECTION)

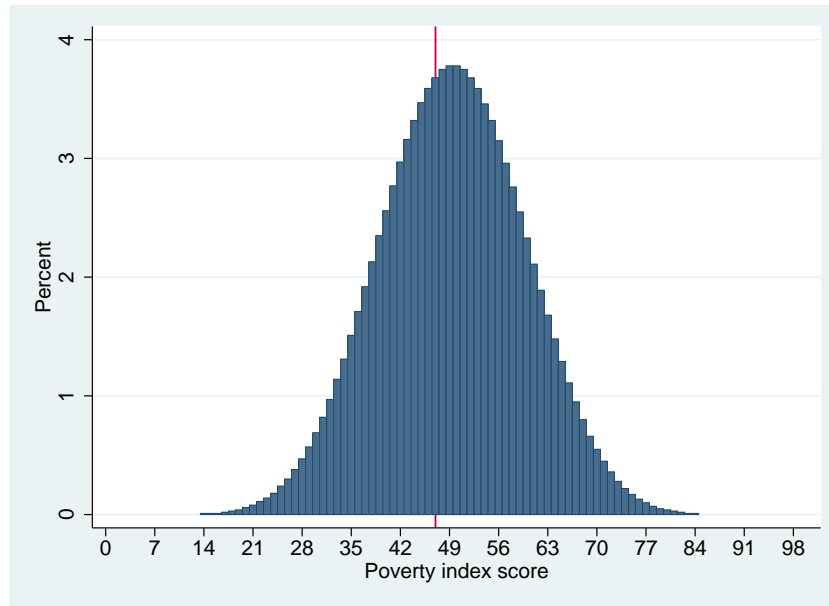
Dependent variable:	discontinuity +/- 3 points				discontinuity +/- 5 points			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of community organizations	-0.027** [0.013]	-0.025* [0.013]			-0.051*** [0.017]	-0.050*** [0.017]		
Newspaper circulation			-0.036** [0.014]	-0.034** [0.014]			-0.065*** [0.021]	-0.063*** [0.021]
Proportion of poor	-0.079 [0.048]	-0.102* [0.052]	-0.126** [0.056]	-0.156** [0.061]	-0.060 [0.056]	-0.070 [0.055]	-0.072 [0.074]	-0.083 [0.070]
Log Population	-0.071 [0.061]	-0.063 [0.063]	-0.052 [0.071]	-0.037 [0.075]	0.026 [0.068]	0.040 [0.071]	0.045 [0.080]	0.071 [0.088]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demography controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography controls		Yes		Yes		Yes		Yes
Observations	669	669	513	513	669	669	513	513
R-squared	0.02	0.02	0.03	0.04	0.02	0.03	0.04	0.05

Note: Robust standard errors in brackets. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%. All regressions include an intercept term and report standardized results. The dependent variable is the difference in the fraction of interviews 3 and 5 points before the threshold relative to the same points after the threshold divided by the number of points, using data for the 6 months prior to the election during election years. The closer to 0 the smaller the discontinuity at the threshold. Demography controls include proportion urban population and proportion poor population using a measure for unsatisfied basic needs from the 1993 and 2005 Population Census. Geography controls include distance to the largest city (capital) in the *departamento* (state) and municipality surface area.

## B. Expansion on Alternative Explanations for Patterns in the Score Distribution

SCORE ALGORITHM. — The Figure below plots the simulated number of possible combinations for each score. The score algorithm takes information from approximately 24 questions. The answers are then used to compute sub-scores for each of the four components. There are 384 possible combinations in the education component, 1008 in the demographic, 90 in utilities, and 480 in the dwelling component for a total of approximately 16 billion possible combinations of answers. The maximum number of combinations is around 600 million for a score of 50. The minimum is 1 for a score of 100. Figure B1 shows that the simulated distribution does not exhibit a discontinuity at the eligibility threshold or anywhere else.<sup>5</sup>

FIGURE B1. SIMULATED DISTRIBUTION OF COMBINATIONS BY SCORE GENERATED BY THE ALGORITHM



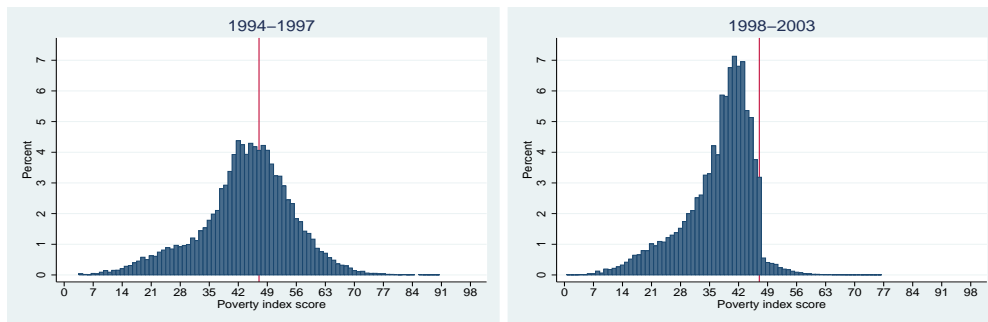
Note: Distribution of number of possible combinations to generate each score. Overall there are approximately 16 billion possible combinations to generate all scores.

<sup>5</sup>Here we assumed that all combinations are equally likely. In reality however, we expect the covariance between certain answers to be different from zero and not to see some combinations in the population. The score distribution depicted in Figure 3 in the main text uses survey data from representative samples of the Colombian population, which we restrict to strata levels below 4 to make it comparable to the population in the Census of the Poor, this distribution does not exhibit discontinuities at the threshold.

SELECTION. — An alternative way to test for selection within a municipality and over time, is to use the geographical information from the 1993 Population Census (prior to the implementation of the Census of the Poor) and see how the characteristics of people in neighborhood blocks interviewed earlier compare to those interviewed later. If the characteristics of people in neighborhood blocks interviewed later are worse than those interviewed earlier then we should expect a left-ward shift in the distribution and be concerned about the possibility of selection.

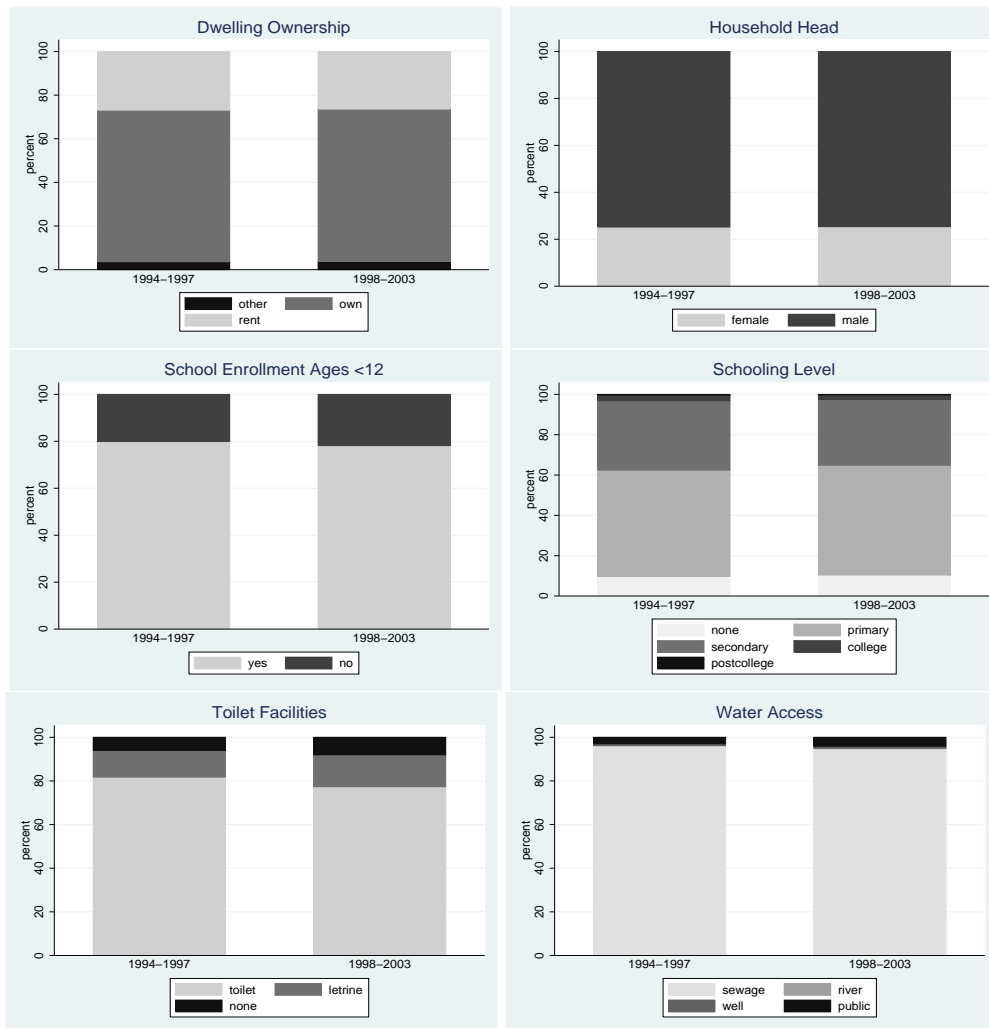
We do this for one municipality, which we call municipality X, with different patterns in the poverty score distribution before and after 1998. Municipality X exhibits a sharp discontinuity at the eligibility threshold in the post-1998 period, see Figure B2. Using the Census of the Poor we identified neighborhood blocks where interviews were conducted before 1998 and blocks where interviews were conducted after 1998. We took this information to the 1993 Population Census to see if the average characteristics of people interviewed after 1998 are worse than those interviewed before 1998. Accounting for the number of households interviewed in each block we graphed the mean characteristics of the people interviewed in the pre and post-1998 period. Overall, the neighborhoods interviewed later do not look worse than the neighborhoods interviewed earlier as seen in Figure B3. Even if this was the case, we would expect a leftward shift in the distribution and not necessarily a sharp discontinuity exactly at the threshold.

FIGURE B2. POVERTY INDEX SCORE DISTRIBUTION PRE AND POST-1998, MUNICIPALITY X



Note: Poverty Index Score distribution for municipality X, for years before (left Figure) and after (right Figure) the score algorithm was released to municipal officials.

FIGURE B3. CHARACTERISTICS OF NEIGHBORHOODS USING THE 1993 POPULATION CENSUS, MUNI. X



### C. Press Articles Documenting Electoral Manipulation in Colombia

“Arrestadas 49 Personas En Comicios” El Tiempo. 20 June 1994. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-155377>. Accessed 7 August 2009.

“Denuncian Posible Fraude Electoral” El Tiempo. 11 March 1994. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-71134>. Accessed 8 August 2009.

“Ramírez, Alcalde Electo De Soacha” El Tiempo. 9 November 1994. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-247901>. Accessed 7 August 2009.

“Demandadas Elecciones De 26 Alcaldes Populares” El Tiempo. 2 December 1994. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-258336>. Accessed 7 August 2009.

“Alcaldes Investigados” El Tiempo. 22 March 1997. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-556887>. Accessed 8 August 2009

“Demandarán Elecciones En Codazzi”. El Tiempo. 4 November 1997. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-667574>. Accessed 8 August 2009

“No cambié Carnets del Sisben por Votos” El Pais. Newspaper. 28 November 1997.

“Irregularidades en la Jornada Electoral” El Tiempo. 4 November 1997. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-694231>. Accessed 7 August 2009.

“Juego de Manos” Semana. 30 March 1998. Magazine online. Available from <http://www.semana.com/noticias-nacion/juegos-manos/37073.aspx>. Accessed 8 August 2009.

“Cómo se Compra un Voto en Colombia.” El Tiempo. 20 June 1998. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-790679#>. Accessed 29 July 2009.

“Políticos Ofrecen Cupos en el Sisben a Cambio de Votos” El Pais. Newspaper. 13 October 2000.

“Procuraduría Alerta en Elecciones” El Tiempo. 27 October 2000. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-1259284>. Accessed 8 August 2009.

“Denuncian Anomalías en las Pasadas Elecciones” El Tiempo. 23 November 2000. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-1231260>. Accessed 8 August 2009.

“La Trampa Electoral” Semana. 8 July 2002. Magazine online. Available from <http://www.semana.com/noticias-nacion/trampa-electoral/1273.aspx>. Accessed 8 August 2009.

“Enriquecimiento a Costa de la Salud” El Tiempo. 30 December 2003. Newspaper online. Available from <http://www.eltiempo.com/archivo/documento/MAM-1045041>. Accessed 8 August 2009.



**D. Poverty Index Score Algorithm**

Description	Weight
<i>Education Component</i>	
Education of the highest wage earner	
Without education	0
Incomplete primary school	1.6239
Complete primary school	3.4435
Incomplete secondary school	5.0039
Complete secondary school	7.3434
Incomplete college	9.7833
Completed college	11.546
Post-graduate	12.4806
Avg. education of household members older than 11 years	
Without education	0
(0, 4]	1.657
(4, 5]	2.9947
(5, 10]	4.969
(10, 11]	7.6387
(11, 15]	9.4425
(15, 16]	10.69
More than 16 of schooling	11.1396
Social security of the highest wage earner	
No social security and self-employed or not working	0
No social security and works in firm of 2-9 workers	1.166
No social security and works in firm of 10 or more workers	2.6545
With social security and self-employed or not working	3.9539
With social security and works in firm of 2-9 workers	5.8427
With social security and works in firm of 10 or more workers	6.9718

Source: Colombia's National Planning Agency, (DNP). Continued on next page.

Description	Weight
<i>Housing Component</i>	
Wall materials	
No walls, bamboo	0
Zinc, cloth, cardboard, metal etc.	0.2473
Unpolished wood	2.0207
Mud	4.8586
Adobe	6.2845
Rock, bricks or blocks	7.7321
Roof materials	
Straw	0
Recycled materials (cardboard, metal, etc)	2.1043
Tiles, zinc (without a ceiling)	3.7779
Tiles, zinc (with a ceiling)	5.0973
Floor materials	
Dirt	0
Unpolished wood	2.9037
Cement	3.6967
Tiles, vinyl or bricks	5.8712
Rugs, polished wood, marble	6.8915
Number of appliances that the household owns	
None	0
Up to 3 basic appliances	2.1435
4 basic appliances without a washer	3.0763
3 to 4 basic appliances with a washer	4.7194

Source: Colombia's National Planning Agency, (DNP). Continued on next page.

Description	Weight
<i>Demographic Component</i>	
Children to family size ratio	
More than 0.65	0
(0.0, .65]	0.2237
No children	1.4761
Employed to family size ratio	
Less than 0.30	0
(0.30, 0.60]	0.6717
(0.60, 0.90]	1.739
More than 0.90	4.0149
Room crowdedness	
Less than 0.20	0
(0.20, 0.30]	0.5584
(0.30, 0.40]	1.6535
(0.40, 0.70]	2.5727
(0.70, 1.00]	4.3886
(1.00, 4.00]	6.0042
More than 4.0	8.3828
Income percapita relative to the minimum wage	
Less than 0.15	0
(0.15, 0.25]	0.8476
(0.25, 0.35]	2.1828
(0.35, 0.50]	3.5362
(0.50, 0.75]	5.3636
(0.75, 1.00]	7.0827
(1.00, 1.25]	8.2489
(1.25, 1.50]	9.4853
(1.50, 2.00]	10.2098
(2.00, 3.00]	11.3999
(3.00, 4.00]	13.0872
More than 4.0	13.7378

Source: Colombia's National Planning Agency, (DNP). Continued on next page.

Description	Weight
<i>Utilities Component</i>	
Water source	
River or spring	0
Public well/pool or other source	1.1601
Well without a pump	2.6497
Well with a pump	4.6037
Truck	6.1693
Water/sewage system	7.2554
Type of toilet facilities	
No toilet facilities	0
Latrine	2.4519
Toilet without connection to water source	3.3323
Toilet connected to a well	3.9615
Toilet connected to sewage	6.8306
Waste collection and disposal	
Throw it to a lot	0
Take it to a container	2.1291
Picked by garbage collection services	3.2701

*Source: Colombia's National Planning Agency, (DNP).*

## REFERENCES

- Besley, Timothy, and Robin Burgess.** 2002. "The Political Economy Of Government Responsiveness: Theory And Evidence From India." *Quarterly Journal of Economics*, 117(4): 1415–1451.
- Fiszbein, Ariel.** 1995. *Local Government Capacity in Colombia: Beyond Technical Assistance*. Washington D.C.: World Bank Country Study.
- Rosas, Andres, and Juan Mendoza.** 2005. "The Economic Effects of Geography: Colombia as a Case Study." Universidad Javeriana Documentos de Economía 003584.
- Schulhofer-Wohl, Sam, and Miguel Garrido.** 2009. "Do Newspapers Matter? Evidence from the Closure of The Cincinnati Post." National Bureau of Economic Research Working Papers 14817.