

Contagious Dishonesty: Corruption Scandals and Supermarket Theft

By GIORGIO GULINO AND FEDERICO MASERA

ONLINE APPENDIX

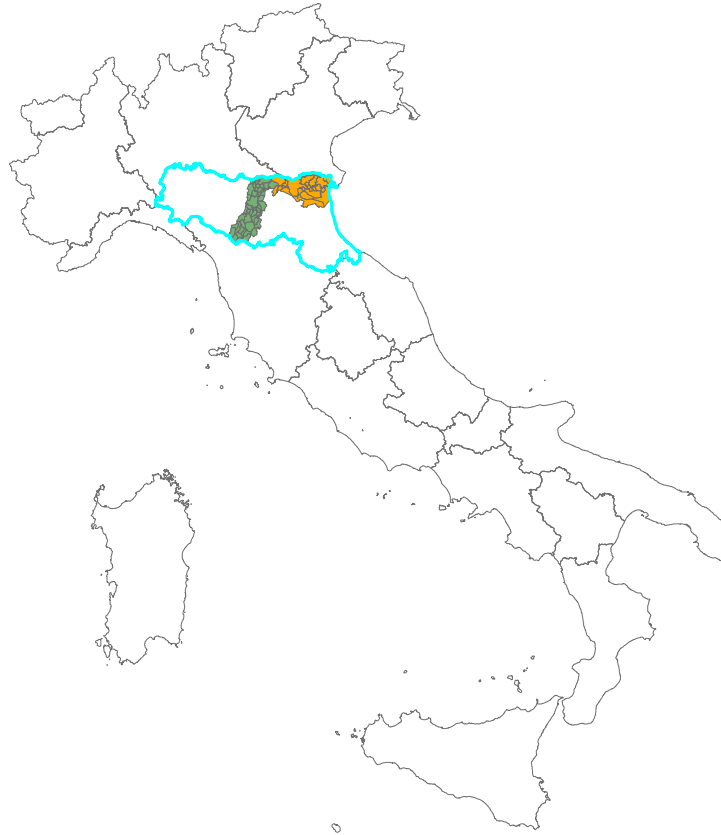


FIGURE A1. PROVINCES OF MODENA AND FERRARA

Note: The maps shows the provinces of Modena (in green) and Ferrara (in orange) in the region Emilia-Romagna (light green edges).

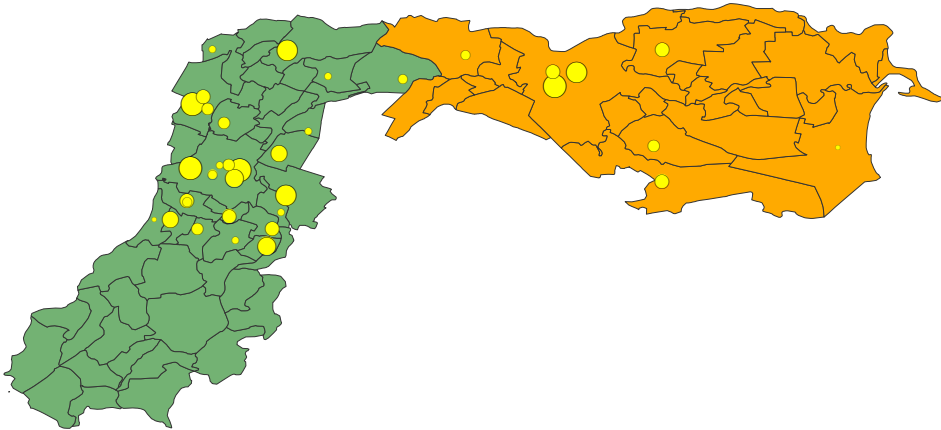


FIGURE A2. SUPERMARKETS *Coop Alleanza 3.0* IN MODENA AND FERRARA

Note: The maps show the municipalities of the provinces of Modena (in green) and Ferrara (in orange). Yellow dots represent the location of supermarkets. The bigger is the dots, the higher is the number of random audits done to shoppers in the supermarket.



(a) Bar-Code Scanner



(b) How to Scan a Product

FIGURE A3. TIME-SAVER TECHNOLOGY: BAR-CODE SCANNER

Note: The figure shows the bar-code scanner that is used by clients that exploit the system called *time-saver technology*.

TABLE A1—NEWS HEADLINE (ITALIAN), MUNICIPALITY AND DAY OF THE TREATMENT

Day	News Headlines	Municipality
23 th November 2016	Abuso edilizio: Bonucchi condannato	Sestola & Serramazzoni
30 th November 2016	Serra, l'ex sindaco resta sotto processo	Serramazzoni
1 st December 2016	I difensori contro l' accusa al processo "pratiche veloci"	Ferrara
3 rd December 2016	In appello mini-condanna per la Modena	Modena
29 th December 2016	Ambrosi-Bonini, la storia dalle origini fino all'usura	Sassuolo, Fiscaglia & Maranello
3 rd February 2017	"Maresciallo, sospensione giusta"	Sassuolo
7 th February 2017	"Appalti e mazzette: processate quei 49"	Carpi
19 th February 2017	"Patto illecito con i privati per devastare il paesaggio"	Serramazzoni
24 th February 2017	Policlinico, dirigente Ccc: "Mai corrotto"	Carpi
20 th April 2017	Soldi e regali in cambio di lavori, arrestati	Palagano
28 th April 2017	La prescrizione devasta il maxi processo	Serramazzoni
19 th May 2017	Corruzione: Rispoli condannato a 5 anni	Carpi & Castelfranco Emilia
9 th June 2017	Illeciti sui funerali? "Ora bisogna indagare"	Ferrara
30 th June 2017	Soldi sottratti a Terre: il buco è di 64mila euro	Argenta
5 th July 2017	"Condannate Baglio e i suoi, Ralenti no"	Serramazzoni
5 th July 2017	Caso Niagara, un'altra condanna	Poggio Renatico
26 th September 2017	Scandalo concorsi truccati. Indagata la Fregni: interdetta dall'ateneo di Modena	Carpi, Modena & Ferrara
4 th October 2017	Caso Cardiologia: oggi attesa la sentenza	Modena
12 th October 2017	"Mazzette in Comune" Un post scatena il sindaco	Castelfranco Emilia
11 th November 2017	Sei anni a Sangiorgi, stangate le aziende	Modena

Note: The table shows the municipality and the day in which the scandal is published on the newspaper and the news headline of the article in Italian. *Source:* the news can be found in on-line archives *Gazzetta di Modena* and *Nuova Ferrara* at the following link, respectively: <https://ricerca.gelocal.it/ricerca/gazzettadimodena> and <https://ricerca.gelocal.it/ricerca/lanuovaferrara>.

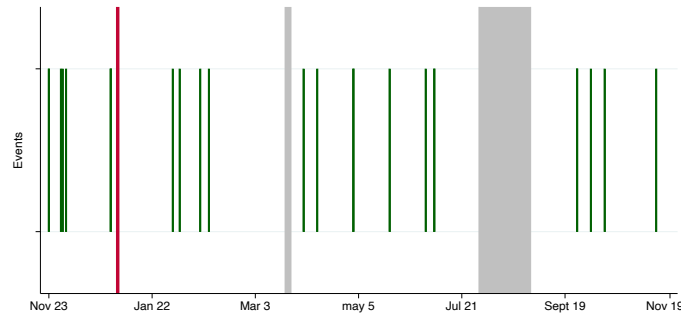


FIGURE A4. DISTRIBUTION OF CORRUPTION SCANDALS OVER THE PERIOD

Note: The figure shows the distribution of corruption scandals over the period of interest. The green vertical lines represent the day of the scandals, the red line represents January 1st, 2017, while the gray areas represent easter break and August 2017, respectively.

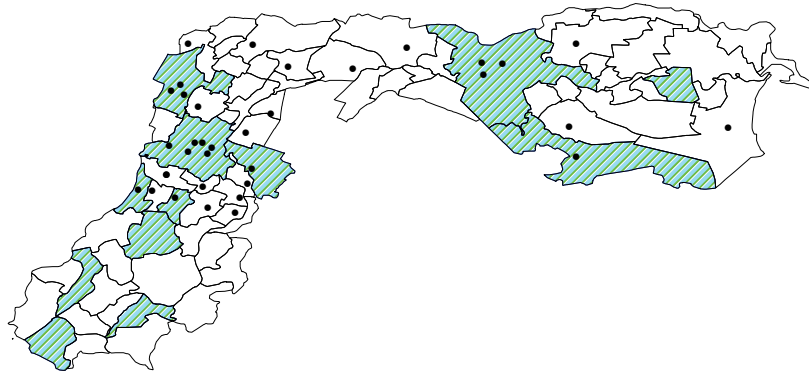


FIGURE A5. DISTRIBUTION OF CORRUPTION SCANDALS ACROSS MUNICIPALITY

Note: The map shows the distribution of scandals across municipalities. Black dots represent the super-market, while the light green municipalities are those in which there is at least a corruption event over the period of interest.

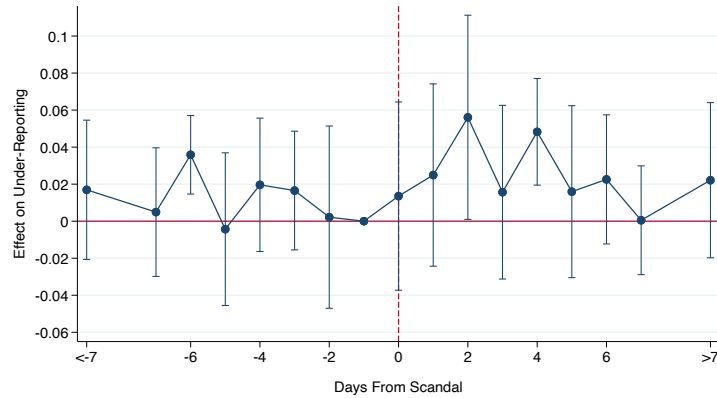


FIGURE A6. CORRUPTION SCANDALS AND UNDER-REPORTING

Note: The graph reports coefficient estimates of the effect of corruption scandal on the probability a customer underreports purchases, using Equation (1) with a window of seven days before and after the news is published. Complete data descriptions and sources are reported in Table A28 in the Appendix, and summary statistics are presented in Table A24.

TABLE A2—CORRUPTION SCANDALS AND UNDER-REPORTING (CONSTANT SAMPLE)

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL	0,0236 (0,0055)	0,0248 (0,0060)	0,0245 (0,0060)	0,0188 (0,0058)	0,0261 (0,0096)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0,14	0,14	0,14	0,14	0,14
Observations	213.857	213.857	213.857	213.857	213.857
R-Square	0,01	0,01	0,01	0,34	0,06

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A3—CORRUPTION SCANDALS AND OVER-REPORTING (CONSTANT SAMPLE)

	OVER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL	-0,0036 (0,0047)	-0,0032 (0,0046)	-0,0028 (0,0047)	-0,0009 (0,0068)	-0,0044 (0,0047)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0,06	0,06	0,06	0,06	0,06
Observations	213.857	213.857	213.857	213.857	213.857
R-Square	0,00	0,01	0,01	0,31	0,06

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to over-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A4—CORRUPTION SCANDALS AND UNDER-REPORTING (ALTERNATIVE CLUSTER)

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL					
Cluster:					
<i>Robust Standard Error</i>	0,022 (0,006)	0,023 (0,006)	0,023 (0,006)	0,018 (0,007)	0,023 (0,010)
<i>Municipality Level × Day Level</i>	0,022 (0,005)	0,023 (0,005)	0,023 (0,005)	0,018 (0,007)	0,023 (0,009)
<i>Municipality × Month Level</i>	0,022 (0,005)	0,023 (0,005)	0,023 (0,005)	0,018 (0,006)	0,023 (0,009)
<i>Shop Level</i>	0,022 (0,005)	0,023 (0,005)	0,023 (0,005)	0,018 (0,007)	0,023 (0,008)
<i>Shop × Day Level</i>	0,022 (0,005)	0,023 (0,005)	0,023 (0,005)	0,018 (0,007)	0,023 (0,010)
<i>Shop × Month Level</i>	0,022 (0,005)	0,023 (0,004)	0,023 (0,004)	0,018 (0,006)	0,023 (0,007)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0,14	0,14	0,14	0,14	0,14
Observations	260.192	260.192	255.749	217.344	255.445
R-Square	0,00	0,01	0,01	0,35	0,05

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard error in the first row, and then clustered at: municipal-day, municipal-month, shop, shop-day and shop-month level are in parentheses.

TABLE A5—CORRUPTION SCANDALS AND OVER-REPORTING (ALTERNATIVE CLUSTER)

	OVER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL					
Cluster:					
<i>Robust Standard Error</i>	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,005)	-0,002 (0,007)
<i>Municipality Level × Day Level</i>	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,007)	-0,002 (0,006)
<i>Municipality × Month Level</i>	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,005)	-0,002 (0,006)
<i>Shop Level</i>	-0,001 (0,003)	-0,001 (0,003)	-0,001 (0,003)	-0,001 (0,005)	-0,002 (0,007)
<i>Shop × Day Level</i>	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,005)	-0,002 (0,007)
<i>Shop × Month Level</i>	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,001 (0,004)	-0,002 (0,006)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Observations	260.192	260.192	255.749	217.344	255.445
R-Square	0,003	0,005	0,006	0,306	0,052

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to over-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard error in the first row, and then clustered at: municipal-day, municipal-month, shop, shop-day and shop-month level are in parentheses.

TABLE A6—CORRUPTION SCANDALS AND UNDER-REPORTING (BOOTSTRAP)

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL					
Bootstrap:					
MUNICIPALITY LEVEL	0.022 (0.009)	0.023 (0.007)	0.023 (0.007)	0.018 (0.009)	0.023 (0.017)
SHOP LEVEL	0.022 (0.005)	0.023 (0.005)	0.023 (0.005)	0.018 (0.007)	0.023 (0.009)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0.14	0.14	0.14	0.14	0.14
Observations	260,192	260,192	255,749	217,344	255,445
R-Square	0.00	0.01	0.01	0.35	0.05

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Bootstrap standard errors clustered at the municipality level and at the shop level in parentheses.

TABLE A7—CORRUPTION SCANDALS AND OVER-REPORTING (BOOTSTRAP)

	OVER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL					
Bootstrap:					
MUNICIPALITY LEVEL	-0.001 (0.006)	-0.001 (0.005)	-0.001 (0.006)	-0.001 (0.007)	-0.002 (0.008)
SHOP LEVEL	-0.001 (0.003)	-0.001 (0.004)	-0.001 (0.003)	-0.001 (0.005)	-0.002 (0.008)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0.06	0.06	0.06	0.06	0.06
Observations	260,192	260,192	255,749	217,344	255,445
R-Square	0.00	0.00	0.01	0.31	0.05

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to over-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Bootstrap standard errors clustered at the municipality level and at the shop level in parentheses.

TABLE A8—BALANCE IN COVARIATES AND OUTCOMES DAYS BEFORE TREATMENT

	Control Group		Treated Group		(1)	(2)	obs.
	mean	SD	mean	SD	diff.	p-value	
Under-Reporting	0,13	0,34	0,14	0,34	-0,01	0,54	40436
Under-Reporting in the Past	0,15	0,36	0,16	0,36	0,00	0,74	8870
Over-Reporting	0,06	0,24	0,06	0,24	0,00	0,79	40436
Total Value Purchases	49,40	46,82	51,77	49,09	0,52	0,61	40436
Number of Products	22,85	20,68	23,59	21,16	0,33	0,29	40436
Morning	0,45	0,50	0,43	0,50	-0,01	0,09	40436
Age	53,64	14,32	54,06	14,58	0,24	0,14	40368
Male	0,40	0,49	0,43	0,49	-0,00	0,63	40436
White-collar Worker	0,31	0,46	0,34	0,48	0,00	0,57	40436
Blue-collar Worker	0,20	0,40	0,18	0,38	0,01	0,19	40436
Retired	0,13	0,34	0,15	0,35	-0,00	0,13	40436
Homemaker	0,08	0,27	0,07	0,25	0,01	0,07	40436
Self Employed	0,05	0,22	0,05	0,22	-0,00	0,44	40436
Business Owner	0,06	0,24	0,05	0,23	-0,00	0,50	40436
Teacher	0,04	0,20	0,05	0,21	-0,00	0,50	40436
Student	0,03	0,17	0,03	0,18	0,00	0,76	40436
Unemployed	0,01	0,11	0,01	0,12	0,00	0,84	40436
Other Employment	0,08	0,27	0,06	0,24	-0,01	0,01	40436
Italian Nationality	0,95	0,21	0,94	0,23	0,00	0,30	40436

Note: Variables description and data sources are reported in Tables A28. For each variable, means and standard deviations in both the control group and the treatment group are reported. Column (1) reports the mean difference between the treatment and the control group; Column (2) reports the p-values of the treatment coefficient of a regression which includes as control municipality and day fixed effect.

TABLE A9—CORRUPTION SCANDALS AND UNDER-REPORTING (ALTERNATIVE SAMPLE)

	BASELINE	EVER SCANDAL	UNDER Vs. OVER	UNDER Vs. ZERO
	(1)	(2)	(3)	(4)
POST SCANDAL	0.023 (0.005)	0.018 (0.005)	0.042 (0.022)	0.025 (0.005)
Municipality FE	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓
Hour of the Day FE	✓	✓	✓	✓
Client Controls	✓	✓	✓	✓
Mean Dependent	0.14	0.14	0.69	0.15
Observations	255,749	149,664	51,298	240,173
R-Square	0.01	0.01	0.02	0.01

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. In column 1 the sample is unrestricted as from Tables 5; in column 2 the sample is restricted cities that have experienced at least a scandal during the period of analysis; in column 3 the sample is restricted to customer that during the sample period have been observed under-reporting or over-reporting purchases at least once; in column 4 we drop from the sample customer that over-report. The *dependent variable* is a dummy taking value 1 if customer is found to over-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A10—CORRUPTION SCANDALS AND UNDER-REPORTING (ALTERNATIVE OUTCOMES)

	OBJECTS STOLEN	VALUE STOLEN (LN)	SHARE VALUE STOLEN
	(1)	(2)	(3)
POST SCANDAL	0,052 (0,021)	0,031 (0,009)	0,180 (0,077)
Municipality FE	✓	✓	✓
Calendar Day FE	✓	✓	✓
Shop FE	✓	✓	✓
Hour of the Day FE	✓	✓	✓
Client Controls	✓	✓	✓
Mean Dependent	1,87	3,55	7,10
Observations	256.189	256.189	256.189
R-Square	0,01	0,01	0,00

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* are: the number of object taken but not declared by the customer while using the time-saver technology (column 1); the total value of products taken but not declared by the customer while using the time-saver technology (column 2); the total value of products taken but not declared by the customer over the total value of purchases (column 3). POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A11—COMPARING TWO-WAY FIXED EFFECTS ESTIMATES WITH SUN, ABRAHAMS (2020)

	GOOGLE TRENDS CORRUZIONE			
	(1)	(2)	(3)	(4)
	Other Regions	Other Topics	Other Crime Topics	Other Years
Two-Way Fixed Effects	5.2137	13.1429	14.4469	8.6577
Sun, Abrahams (2020)	5.7068	14.5231	10.2997	10.2997
P-value Difference	0.7982	0.6119	0.5680	0.5680

Note: This table report the classical two-way fixed effects estimate in the first row and the average treatment effects estimated following Sun and Abraham 2020 method in the second row. The last row displays the p-value of test that has a null hypothesis that the two-way fixed effects estimate is equal to the average treatment effects estimated following Sun and Abraham 2020.

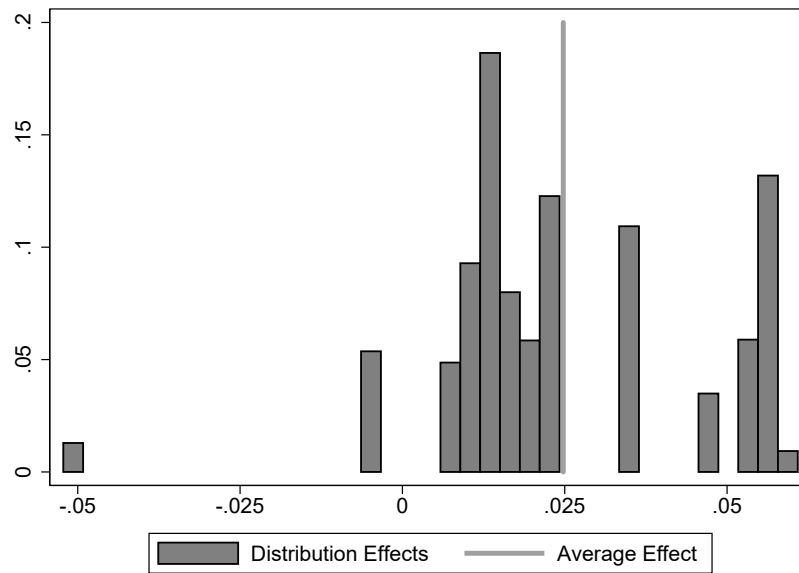


FIGURE A7. DISTRIBUTION OF THE TREATMENT EFFECTS

Note: The graph reports distribution of the effect of each corruption scandal on the likelihood of under-reporting. The vertical line shows the average of these effects.

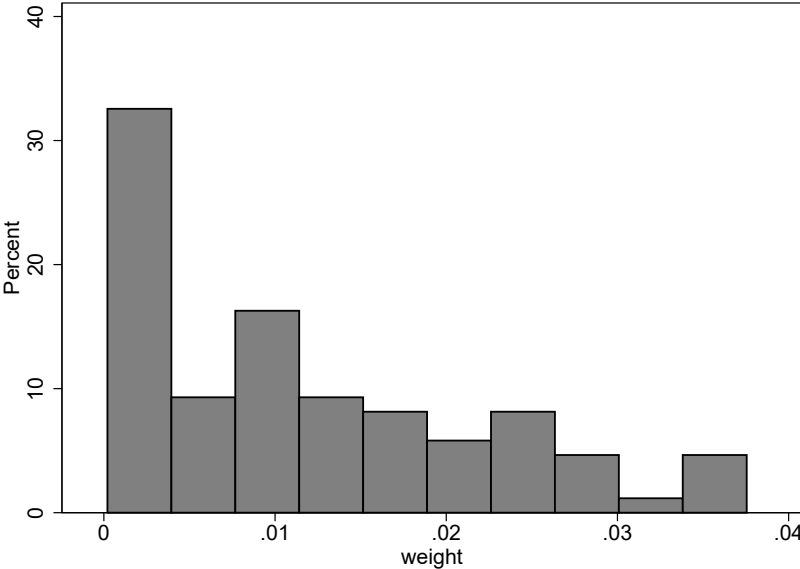


FIGURE A8. DISTRIBUTION OF THE WEIGHTS OF THE TREATMENT EFFECT

Note: De Chaisemartin and d’Haultfoeuille 2020 shows that in a two-way fixed effect setting the estimated treatment effect is a weighted average of each treatment effect. The figure reports distribution of these weights.

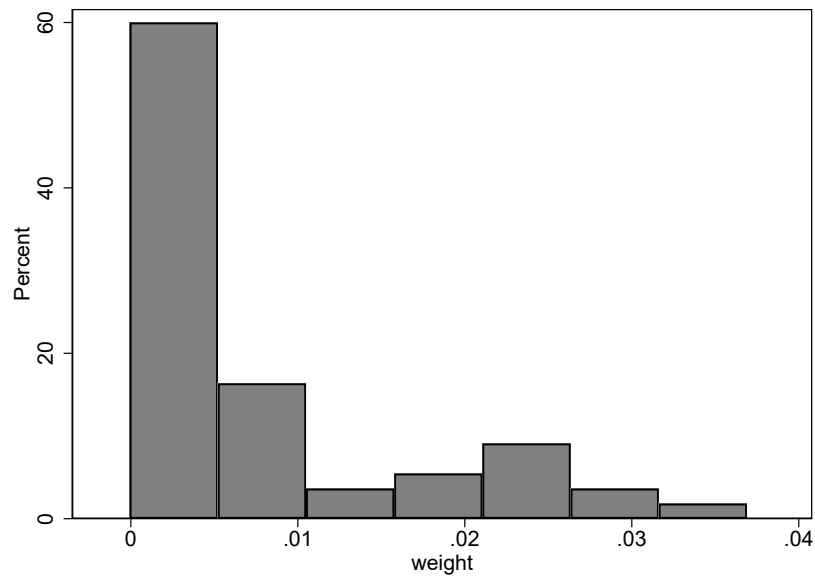


FIGURE A9. DISTRIBUTION OF THE WEIGHTS OF THE GOOGLE NEWS EFFECT

Note: De Chaisemartin and d'Haultfoeuille 2020 shows that in a two-way fixed effect setting the estimated treatment effect is a weighted average of each treatment effect. The figure reports distribution of these weights.

TABLE A12—CONTROLLING FOR THE CLIENT PAST BEHAVIOR

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL	0,0288 (0,0029)	0,0306 (0,0034)	0,0314 (0,0031)	0,0291 (0,0052)	0,0254 (0,0105)
Municipality FE	✓	✓	✓	×	✓
Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0,14	0,14	0,14	0,14	0,14
Observations	157.106	157.106	155.120	130.971	154.314
R-Square	0,02	0,02	0,02	0,36	0,09

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A13—SAME SAMPLE AS A12 WITHOUT THE CONTROL

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL	0,0300 (0,0027)	0,0318 (0,0032)	0,0324 (0,0032)	0,0287 (0,0057)	0,0285 (0,0107)
Municipality FE	✓	✓	✓	×	✓
Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Mean Dependent	0,14	0,14	0,14	0,14	0,14
Observations	157.106	157.106	155.120	130.971	154.314
R-Square	0,01	0,01	0,01	0,33	0,08

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A14—GOOGLE TRENDS ABOUT THE WORD “CORRUPTION”

Dep. Var. GOOGLE TRENDS ABOUT THE WORD CORRUPTION	
	(1)
POST SCANDAL	14,642 (3,560)
Municipality FE	✓
Observations	260.651
R-Square	0,01

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is the number of searches for the word corruption in the Region Emilia Romagna. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A15—GOOGLE TRENDS ABOUT THE WORD “CORRUPTION” VS. OTHER CONTROL TERMS

	GOOGLE TRENDS CORRUPTION			
	(1)	(2)	(3)	(4)
Post Scandal \times Treated Index	0.2357 (0.0800)	0.6272 (0.0737)	0.7048 (0.0579)	0.1165 (0.0630)
Day FE	✓	✓	✓	✓
Search Term FE	✓	✓	✓	✓
Control Terms	Other Regions	Other Topics	Other Crime Topics	Other Years
Observations	4,602	2,478	2,124	1,770
R-Square	0.34	0.62	0.54	0.41

Note: OLS estimates. Observations are at the Google index/Day level. The *dependent variable* is the Google trends index score. In all columns the *Treated index* is the Google Trends index for the word corruption during the year of our study in Emilia-Romagna. In Column (1) the control indexes are the Google Trends index for the word corruption during the year of our study in other Italian regions. In Column (2) the control indexes are the Google Trends index for the topics “football”, “restaurants”, “movies”, “online shopping”, “job opportunities” and “travel” during the year of our study in Emilia-Romagna. In Column (3) the control indexes are the Google Trends index for robberies, homicides, shootings, burglaries and rape during the year of our study in Emilia-Romagna. In Column (4) the control indexes are Google Trends indexes for the word corruption in Emilia-Romagna in the two years prior and the two years after the time period of our study. *Post Scandal* is the share of people that had a corruption scandal in their municipality in the last four days. Regressions are weighted by the number of people we observe in our dataset. Robust standard errors clustered at the Google Trends index level are in parentheses.

Figure A10 shows the search activity for some scandals. The blue lines show the normalized scores reported by Google Trends while the gray areas identify the week of the scandal. For example, the scandal of Maria Cecilia Fregni (graph (a) of Figure A10) bursts for the first time during the period of analysis. She is an university professor of law, who was involved in a big scandal of “competition rigging” in public universities. Another important case is the so called “Operazione Teseo” (graph (b) in Figure A10) about the ex-mayor of Seramazzoni, municipality in the province of Modena, who was involved in a case of rigged contracts discovered for the first time in 2012, that over the years has had many important developments, involving new administrators and bureaucrats. Indeed, for this keyword the graph shows many peaks both before and after the week identified in the newspapers. Also the scandal about Domenico Guigli, ex-mayor of Palagano, a small municipality in the province of Modena, is a case that broke out in 2003, and had some recent developments. However, for this scandal we do not find any other peak of search activities in the reference period, probably due to the fact that it is an old scandal or that Palagano is a small municipality.

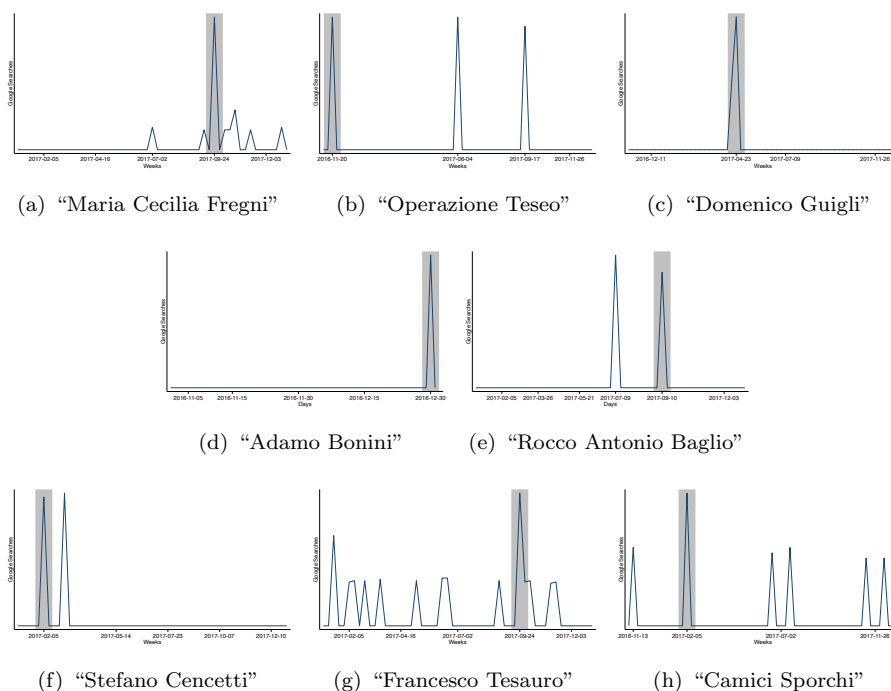


FIGURE A10. GOOGLE TRENDS SEARCH ACTIVITY ABOUT SCANDALS

Note: The graphs report Google trends search activity about the names of public officials involved into corruption scandals. Table A1, in Appendix, shows the news headline of the articles (in Italian).

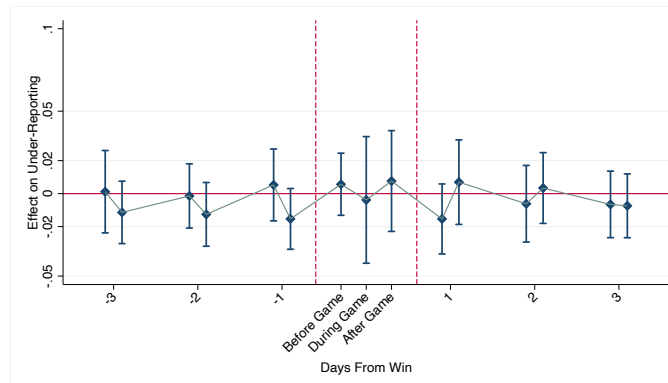


FIGURE A11. WINNING A FOOTBALL GAME ON UNDER-REPORTING

Note: The graph reports coefficient estimates of the effect of winning a football game (conditional on having a game) on the probability a customer under-report purchases, in the spirit of equation (1), using a window of three days before and after the day of the game. For each day there are two bins: morning (befor 14pm, included) and afternoon (after 14pm). As for game day at 0 there are three bins, before, during and after the football game. Robust standard errors are bootstrapped. Complete data descriptions, data sources are reported in Table A28 in Appendix, and summary statistics are presented in Table A26.

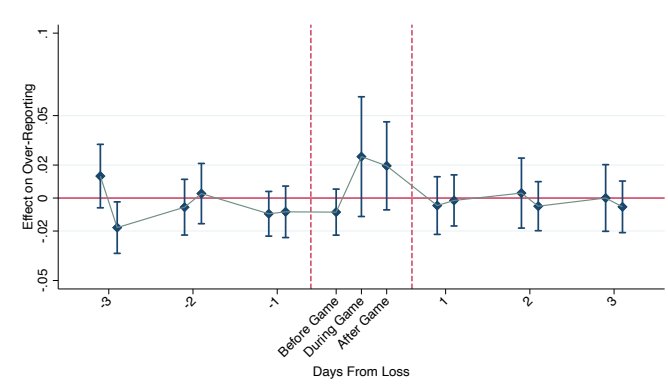


FIGURE A12. LOSING A FOOTBALL GAME ON OVER-REPORTING

Note: The graph reports coefficient estimates of the effect of losing a football game (conditional on having a game) on the probability a customer over-report purchases, in the spirit of equation (1), using a window of three days before and after the day of the game. For each day there are two bins: morning (befor 14pm, included) and afternoon (after 14pm). As for game day at 0 there are three bins, before, during and after the football game. Robust standard errors are bootstrapped. Complete data descriptions, data sources are reported in Table A28 in Appendix, and summary statistics are presented in Table A26.

TABLE A16—NEWS CATEGORIES

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST ANY NEWS	0,0080 (0,0025)				
POST SCANDAL		0,0234 (0,0052)			0,0254 (0,0047)
POST POSITIVE			0,0030 (0,0042)		0,0043 (0,0056)
POST NEUTRAL				-0,0009 (0,0026)	-0,0023 (0,0023)
Municipality FE	✓	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓	✓
Hour of the Day FE	✓	✓	✓	✓	✓
Client Controls	✓	✓	✓	✓	✓
Mean Dependent					
Observations	255.749	255.749	255.749	255.749	255.749
R-Square	0,01	0,01	0,01	0,01	0,01

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST ANY NEWS is a dummy variable that takes value 1 if a given day is in the first four days after any news with the word corruption is made public in the municipality of the client and zero otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. POST POSITIVE is a dummy variable that takes value 1 if a given day is in the first four days after news with positive information is made public in the municipality of the client and zero otherwise. POST HIGHER COURT is a dummy variable that takes value 1 if a given day is in the first four days after news about a higher court decisions is made public in the municipality of the client and zero otherwise. POST OTHER NEUTRAL is a dummy variable that takes value 1 if a given day is in the first four days after news with neutral information about corruption case is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A17—GOOGLE TRENDS ABOUT THE WORD “CORRUPTION” VS. OTHER CONTROL TERMS

	GOOGLE TRENDS CORRUPTION			
	(1)	(2)	(3)	(4)
POST POSITIVE × TREATED INDEX	-0.9385 (2.1456)	1.3823 (2.0875)	-0.6217 (3.0360)	4.1168 (2.9122)
POST NEUTRAL × TREATED INDEX	2.7587 (2.1640)	1.2451 (0.6826)	1.9858 (1.3290)	1.1798 (1.4745)
POST SCANDAL × TREATED INDEX	6.2253 (2.0250)	14.8748 (2.8854)	16.2852 (3.3227)	9.6653 (3.1509)
Calendar Day FE	✓	✓	✓	✓
Index FE	✓	✓	✓	✓
Control Index	Other Regions	Other Topics	Other Crime Topics	Other Years
Observations	3,388,502	1,824,578	1,563,924	1,303,270
R-Square	0.34	0.62	0.53	0.29

Note: OLS estimates. The *dependent variable* is the Google trends index score. In all columns the *Treated index* is the Google Trends index for the word corruption during the year of our study in Emilia-Romagna. In Column (1) the control indexes are the Google Trends index for the word corruption during the year of our study in other Italian regions. In Column (2) the control indexes are the Google Trends index for the topics “football”, “restaurants”, “movies”, “online shopping”, “job opportunities” and “travel” during the year of our study in Emilia-Romagna. In Column (3) the control indexes are the Google Trends index for robberies, homicides, shootings, burglaries and rape during the year of our study in Emilia-Romagna. In Column (4) the control indexes are Google Trends indexes for the word corruption in Emilia-Romagna in the two years prior and the two years after the time period of our study. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. POST POSITIVE is a dummy variable that takes value 1 if a given day is in the first four days after news with positive information is made public in the municipality of the client and zero otherwise. POST NEUTRAL is a dummy variable that takes value 1 if a given day is in the first four days after news with neutral information about corruption case is made public in the municipality of the client and zero otherwise. Robust standard errors clustered at the municipality and Google Trends index level are in parentheses.

TABLE A18—CORRUPTION SCANDALS AND SELECTION INTO TREATMENT

	AGE	MALE	WHITE-C. WORKER	BLUE-C. WORKER	RETIRED	HOMEMAKER	SELF-EMPLOYED	B. OWNER
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
POST SCANDAL	0.056 (0.148)	0.003 (0.003)	0.005 (0.007)	-0.009 (0.004)	0.001 (0.002)	0.008 (0.002)	-0.003 (0.002)	-0.001 (0.003)
Mean Dependent	53.52	0.41	0.32	0.21	0.13	0.08	0.05	0.06
Observations	260,166	260,651	257,416	257,416	257,416	257,416	257,416	257,416
R-Square	0.03	0.02	0.01	0.02	0.02	0.01	0.01	0.00

	TEACHER	STUDENT	UNEMPLOYED	OTHER E.	ITALIAN	MORNING	EVENING	N. CLIENTS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
POST SCANDAL	0.003 (0.002)	-0.001 (0.003)	0.003 (0.001)	-0.005 (0.002)	0.009 (0.003)	0.012 (0.004)	-0.002 (0.003)	-1.388 (1.926)
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dependent	0.04	0.03	0.01	0.08	0.94	0.27	0.04	15.43
Observations	257,416	257,416	257,416	257,416	260,651	260,651	260,651	16,885
R-Square	0.01	0.01	0.00	0.03	0.02	0.04	0.03	0.92

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. *dependent variable* are: the age of the customer (column 1.a); a dummy taking value 1 if customer is male and 0 otherwise (column 2.a); a dummy taking value 1 if customer is a white collar-employee and 0 otherwise (column 3.a); a dummy taking value 1 if customer is a blue-collar worker and 0 otherwise (column 4.a); a dummy taking value 1 if customer is retired and 0 otherwise (column 5.a); a dummy taking value 1 if customer is a housewife and 0 otherwise (column 6.a); a dummy taking value 1 if customer is self-employed and 0 otherwise (column 7.a); a dummy taking value 1 if customer is a business owner and 0 otherwise (column 8.a); a dummy taking value 1 if customer is a teacher and 0 otherwise (column 1.b); a dummy taking value 1 if customer is a student and 0 otherwise (column 2.b); a dummy taking value 1 if customer is unemployed and 0 otherwise (column 3.a); a dummy taking value 1 if customer employment is not classified under any of the previous category and 0 otherwise (column 4.a); a dummy taking value 1 if customer has Italian nationality and 0 otherwise (column 5.b); a dummy taking value 1 if the audit was done in the morning, before 12am and zero otherwise (column 6.b); a dummy taking value 1 if the audit was done in the evening, after 6pm and zero otherwise (column 7.b); the total number of customers that go to the supermarket (column 8.b). POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A19—CORRUPTION SCANDALS AND PURCHASING BEHAVIOR

	TOTAL VALUE	TOTAL VALUE	N. OBJECTS	N. OBJECTS	AVG VALUE	AVG VALUE
	(1)	(2)	(3)	(4)	OBJECTS	OBJECTS
POST SCANDAL	0.951 (0.363)	-0.819 (1.722)	0.259 (0.198)	-0.536 (0.816)	-0.043 (0.048)	0.105 (0.043)
Municipality FE	✓	✓	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓	✓	✓
Hour of the Day FE	✓	✓	✓	✓	✓	✓
Client Controls	✓	✓	✓	✓	✓	✓
Mean Dependent	51.74	69.80	23.80	32.65	2.62	2.32
Observations	255,749	35,271	255,749	35,271	255,693	35,271
R-Square	0.12	0.15	0.08	0.11	0.02	0.05

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. In column 2, 4 and 6 the sample is restricted to customer that during the sample period have been observed under-reporting purchases at least once. *dependent variable* are: the total value of object purchased (columns 1 & 2); the total number of products purchased (columns 3 & 4); the average value of a products (columns 5 & 6). POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A20—CORRUPTION SCANDALS AND UNDER-REPORTING - NEIGHBOURING MUNICIPALITIES

	UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST SCANDAL	0,022 (0,005)	0,024 (0,005)	0,024 (0,005)	0,017 (0,005)	0,027 (0,008)
POST SCANDAL NEIGHBORING M.	0,000 (0,005)	0,002 (0,004)	0,002 (0,004)	-0,001 (0,005)	0,006 (0,007)
Municipality FE	✓	✓	✓	×	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	×	✓	✓	✓	✓
Hour of the Day FE	×	✓	✓	✓	✓
Client Controls	×	×	✓	×	✓
Client FE	×	×	×	✓	×
Shop FE × Day FE	×	×	×	×	✓
Observations	260.192	260.192	255.749	217.344	255.445
R-Square	0,005	0,008	0,010	0,345	0,055

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. POST SCANDAL NEIGHBORING M. is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the neighboring municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A21—THE MORAL COST OF STEALING (SUB-GROUPS)

	Dep. Var. UNDER-REPORTING						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
POST SCANDAL	0,024 (0,005)	0,027 (0,005)	0,028 (0,006)	0,017 (0,005)	0,020 (0,004)	0,024 (0,006)	0,007 (0,005)
POST SCANDAL × VAR. H	-0,025 (0,008)	-0,055 (0,012)	-0,029 (0,006)	0,018 (0,005)	0,018 (0,015)	-0,010 (0,019)	0,029 (0,010)
TOTAL EFFECT	-0,001 (0,010)	-0,028 (0,015)	-0,002 (0,005)	0,035 (0,007)	0,038 (0,016)	0,014 (0,018)	0,036 (0,009)
Municipality FE	✓	✓	✓	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓	✓	✓	✓
Hour of the Day FE	✓	✓	✓	✓	✓	✓	✓
Client Controls	✓	✓	✓	✓	✓	✓	✓
VAR. H	Client is a Student	Client is a Homemaker	Client is Retired	Client is an Employee	Client is a Process Worker	Client is a Rich Taxpayer	Client is a Poor Taxpayer
Mean Dependent	0,14	0,14	0,14	0,14	0,14	0,14	0,14
Observations	255.749	255.749	255.749	255.749	255.749	255.749	255.749
R-Square	0,01	0,01	0,01	0,01	0,01	0,01	0,01

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST SCANDAL is a dummy variable that takes value 1 if a given day is in the first four days after a corruption scandal is made public in the municipality of the client and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A24. Robust standard errors clustered at the municipality level are in parentheses.

TABLE A22—LOSING A FOOTBALL GAME ON UNDER-REPORTING

	Dep. Var. UNDER-REPORTING				
	(1)	(2)	(3)	(4)	(5)
POST LOSS	0.028 (0.014)	0.028 (0.019)	0.025 (0.015)	0.035 (0.024)	0.020 (0.024)
POST LOSS × VAR. H		0.001 (0.024)	0.013 (0.027)	-0.011 (0.026)	0.012 (0.027)
TOTAL EFFECT	0.028 (0.013)	0.029 (0.017)	0.038 (0.028)	0.024 (0.016)	0.033 (0.017)
Municipality FE	✓	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓	✓
Hour of the Day FE	✓	✓	✓	✓	✓
Client Controls	✓	✓	✓	✓	✓
VAR. H		Client is Female	Client is not Taxpayer	Client is Taxpayer	Client is Working Age
Mean Dependent	0.14	0.14	0.14	0.14	0.14
Observations	65,122	65,122	65,122	65,122	65,122
R-Square	0.01	0.01	0.01	0.01	0.01

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST LOSS is a dummy variable that takes value 1 if a given hour is in the first twenty four hours after the football team of client municipality lost the game and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A26. Bootstrapped robust standard errors are in parentheses.

TABLE A23—LOSING A FOOTBALL GAME ON UNDER-REPORTING

Dep. Var. UNDER-REPORTING							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
POST LOSS	0.028 (0.014)	0.031 (0.013)	0.022 (0.015)	0.025 (0.015)	0.023 (0.016)	0.035 (0.012)	0.015 (0.020)
POST LOSS × VAR. H	0.027 (0.071)	-0.037 (0.048)	0.038 (0.043)	0.010 (0.026)	0.029 (0.033)	-0.084 (0.035)	0.025 (0.027)
TOTAL EFFECT	0.054 (0.071)	-0.006 (0.043)	0.060 (0.034)	0.035 (0.022)	0.053 (0.033)	-0.049 (0.035)	0.040 (0.019)
Municipality FE	✓	✓	✓	✓	✓	✓	✓
Calendar Day FE	✓	✓	✓	✓	✓	✓	✓
Shop FE	✓	✓	✓	✓	✓	✓	✓
Hour of the Day FE	✓	✓	✓	✓	✓	✓	✓
Client Controls	✓	✓	✓	✓	✓	✓	✓
VAR. H	Client is a Student	Client is a Homemaker	Client is Retired	Client is an Employee	Client is a Process Worker	Client is a Rich Taxpayer	Client is a Poor Taxpayer
Mean Dependent	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Observations	65,122	65,122	65,122	65,122	65,122	65,122	65,122
R-Square	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Note: OLS estimates. The unit of observation is the customer, resident in a given municipality and audited a given day. The *dependent variable* is a dummy taking value 1 if customer is found to under-report at least a product while shopping and 0 otherwise. POST LOSS is a dummy variable that takes value 1 if a given hour is in the first twenty four hours after the football team of client municipality lost the game and zero otherwise. Complete data descriptions and data sources are presented in Table A28 in Appendix, while summary statistics are presented in Table A26. Bootstrapped robust standard errors are in parentheses.

TABLE A24—SUMMARY STATISTICS

	N	Mean	Std. Dev.	Min.	Max.
Total Obs	260195
Clients	103036	2.5	2.5	1	73
Shops	35	7447	7592	698	35866
Municipalities	78	3342	8356	1	59528
Customer Characteristics:					
Age	.	54	14	18	109
Male	.	41	49	0	100
White collar employee	.	32	47	0	100
Blue collar employee	.	21	40	0	100
Retired	.	13	34	0	100
Housewife	.	7.7	27	0	100
Self employed	.	4.9	22	0	100
Business owner	.	6.1	24	0	100
Teacher	.	4.1	20	0	100
Student	.	3	17	0	100
Unemployed	.	1.3	12	0	100
Other employment	.	7.6	26	0	100
Italian nationality	.	94	24	0	100
Province of Birth	116	2247	12340	5	123959
Audits Record:					
Under-Reporting	.	14	35	0	100
Over-reporting	.	6.1	24	0	100
Total Value	.	52	48	0	709
Share of Value of Under-Reported	.	7.1	11	.004	179
Share of Value of Over-Reported	.	6.6	15	.004	200

TABLE A25—SUMMARY STATISTICS FOR AUDITS DONE TO NON-RESIDENTS

	N	Mean	Std. Dev.	Min.	Max.
Total Obs	21610
Clients	11893	1.8	1.6	1	35
Shops	35	619	1136	13	4912
Municipalities	1311	17	89	1	1810
Customer Characteristics:					
Age	.	46	14	18	102
Male	.	49	50	0	100
White collar employee	.	31	46	0	100
Blue collar employee	.	18	39	0	100
Retired	.	5.5	23	0	100
Housewife	.	5.2	22	0	100
Self employed	.	5.2	22	0	100
Business owner	.	6.7	25	0	100
Teacher	.	2.7	16	0	100
Student	.	9.6	29	0	100
Unemployed	.	1.5	12	0	100
Other employment	.	14	35	0	100
Italian nationality	.	93	25	0	100
Province of Birth	114	190	506	1	3346
Audits Record:					
Under-Reporting	.	17	38	0	100
Over-reporting	.	6	24	0	100
Total Value	.	63	57	0	897
Share of Value of Under-Reported	.	6.5	10	.0048	179
Share of Value of Over-Reported	.	7.1	17	.013	175

TABLE A26—SUMMARY STATISTICS FOR AUDITS DONE TO NON-RESIDENTS

	N	Mean	Std. Dev.	Min.	Max.
Total Obs	125583
Clients	50343	2.5	2.4	1	63
Shops	35	3595	6293	8	27052
Municipalities	4	31456	22030	6508	59528
Customer Characteristics:					
Age	.	54	15	18	107
Male	.	44	50	0	100
White collar employee	.	34	47	0	100
Blue collar employee	.	17	38	0	100
Retired	.	14	35	0	100
Housewife	.	6.8	25	0	100
Self employed	.	5.9	23	0	100
Business owner	.	5.8	23	0	100
Teacher	.	4.8	21	0	100
Student	.	3.5	18	0	100
Unemployed	.	1.3	11	0	100
Other employment	.	6.3	24	0	100
Italian nationality	.	93	25	0	100
Province of Birth	116	1085	5558	5	54004
Audits Record:					
Under-Reporting	.	14	35	0	100
Over-reporting	.	6.1	24	0	100
Total Value	.	53	49	0	709
Share of Value of Under-Reported	.	7.2	11	.004	166
Share of Value of Over-Reported	.	6.9	16	.0055	200
Games Record:					
Loosing a Game	.	.038	.19	0	1
Winning a Game	.	.024	.15	0	1

TABLE A27—SUMMARY STATISTICS (WHOLE POPULATION)

	Mean Population	Mean Our Sample
Province of Modena:		
Age ¹	43.4	52.9
Male ¹	48.9	40.3
Retired Men ²	21.6	16.0
Retired Women ²	22.3	11.1
All types of self-employed ^{3 *}	25.8	15.5
All types of employee ^{3 **}	74.2	84.4
Italian nationality ⁴	90.3	93.3
Province of Ferrara:		
Age ¹	47.4	55.3
Male ¹	48.1	43.0
Retired Men ²	26.4	17.1
All types of self-employed ^{3 *}	21.6	16.2
All types of employee ^{3 **}	78.4	83.8
Retired Women ²	26.1	8.7
Italian nationality ³	93.6	96.1

Source (1): year 2019

Source (2): year 2019

Source (3): year 2016

Source (4): year 2016

* Categories included: Self employed and Business owner

** Categories included: White collar employee, Blue collar employee, Teacher

TABLE A28—VARIABLES DESCRIPTION AND DATA SOURCES: MAIN VARIABLES

Main treatment:

Post scandal. It is a dummy taking value 1 the four day after is published a news about a scandal involving a public official that works in municipality of the customer and 0 otherwise

Main outcomes:

Under-reporting. It is a dummy taking value 1 if customer is found to under-report at least a product and 0 otherwise.

Over-reporting. It is a dummy taking value 1 if customer is found to over-report at least a product and 0 otherwise.

Alternative outcomes:

Object stolen. Number of object taken but not declared by the customer while using the time-saver technology.

Value stolen (ln). Total value of products taken but not declared by the customer while using the time-saver technology.

Share value stolen. Total value of products taken but not declared by the customer over the total value of purchases.

Customer characteristics:

Age. The age of the customer.

Male. It is a dummy taking value 1 if customer is male and 0 otherwise.

White collar employee. It is a dummy taking value 1 if customer is employee and 0 otherwise.

Blue collar employee. It is a dummy taking value 1 if customer is a process worker and 0 otherwise.

Retired. It is a dummy taking value 1 if customer is retired and 0 otherwise.

Housewife. It is a dummy taking value 1 if customer is a housewife and 0 otherwise.

Self-employed. It is a dummy taking value 1 if customer is self-employed and 0 otherwise.

Business owner. It is a dummy taking value 1 if customer is a business owner and 0 otherwise.

Teacher. It is a dummy taking value 1 if customer is a teacher and 0 otherwise.

Student. It is a dummy taking value 1 if customer is a student and 0 otherwise.

Unemployed. It is a dummy taking value 1 if customer is unemployed and 0 otherwise.

Other employment. It is a dummy taking value 1 if customer employment is not classified under any of the previous category and 0 otherwise.

Italian nationality. It is a dummy taking value 1 if customer has Italian nationality and 0 otherwise.

Province of birth. The province of birth of the customer.

TABLE A29—VARIABLES DESCRIPTION AND DATA SOURCES: MAIN VARIABLES

Other characteristics:

Morning. It is a dummy taking value 1 if the audit was done in the morning, before 12am and zero otherwise.

Evening. It is a dummy taking value 1 if the audit was done in the evening, after 6pm and zero otherwise.

Number of Clients. The total number of customers that go to the supermarket.

Purchasing behavior:

Total value. The total value of object purchased.

Total number of objects. The total number of products purchased.

Average value objects. The average value of a product.

Interaction terms:

Article number of words. The number of words of the newspaper article which we exploit to identify the case of corruption scandals.

First Pages of Newspaper. It is a dummy variable equal to one if the news is published within the first seven pages and zero otherwise.

Match-Day. It is a dummy variable equal to one if the day in which the news is published has been played a game of the football team of the municipality and zero otherwise.

Match-Day or Day After. It is a dummy variable equal to one if the day, of the four days after, in which the news is published has been played a game of the football team of the municipality and zero otherwise.

Small supermarket. It is a dummy taking values one weather the supermarket has less than the median number of clients within a year and zero otherwise.

Supermarket in Small Municipality. It is a dummy taking values one weather the supermarket of the shopping visit is in a municipality that has less than the 30,000 inhabitants and zero otherwise.

Crowded hours. It is a dummy taking values one weather within the reference hour of the day there are less than the median number of clients within a hour and zero otherwise.

Home supermarket. It is a dummy taking values if the client go shopping in a supermarket located in the municipality of residence and zero otherwise.

In-group gender. It is a dummy taking values if the client and the public official involved in the corruption scandal are of the same gender and zero otherwise.

Client is not Taxpayer. It is a dummy taking values if the client self-report of being housewife, student, unemployed or retired and zero otherwise

Client is Taxpayer. It is a dummy taking values if the client self-report of being white collar employee, blue collar employee, self-employed, business owner, teacher and zero otherwise.

Working Age. It is a dummy taking values if the client of age between 28 and 65 and zero otherwise.