Benefits by luck: A study of lotteries as a selection method for government programs

— Online Appendix[1] —

Table of Contents

A	Support for Public Policy Lotteries in the Scholarly Literature	2
B	Housing Lotteries Rules	2
C	Survey Sampling	3
D	Summary Statistics — Quasi-Experimental Analysis	5
E	Original Survey Question Wording	7
F	Balance Tests — Quasi-Experimental Analysis	8
G	Attrition — Quasi-Experimental Analysis	14
H	Outcome Variable Descriptives — Quasi-Experimental Analysis	18
I	Tables of Results — Quasi-Experimental Analysis	19
J	Additional Quasi-Experimental Analyses	23
K	Complier Average Treatment Effects	25
L	Qualitative Semi-Structured Interviews	28
M	Design of Survey Experiment & Survey Sampling	30
N	Balance Tests – Experimental Data	33
O	Additional Analysis – Experimental Data	33
P	Considerations about the Pre-analysis Plans	40
Q	Ethics Discussion	41
R	Data Sources & Transparency	42

[1]Parts of sections B–G and J–L of this online appendix may overlap or partly reproduce background information from other manuscripts currently under review or in preparation that use the same datasets.

Appendix A Support for Public Policy Lotteries in the Scholarly Literature

We searched for the words "lottery," "lotteries", "by lot," and "sortition" in titles of papers published since the year 2000 in top journals in Political Science (American Political Science Review, American Journal of Political Science, Journal of Politics, and Comparative Political Studies), Public Administration (Public Administration Review, Journal of Public Administration Research & Theory, Governance, and Public Administration), Economics (American Economic Review, Journal of Political Economy, Quarterly Journal of Economics, Review of Economic Studies), and Development Studies (World Development, Development and Change, Journal of Development Studies, Development Policy Review).

The 30 papers that were returned referred to studies about state lotteries in the US, random audits, measuring the impact of the Vietnam lotteries, or, more commonly, studies that used lotteries as part of their design.

The single papers published since 2000 in leading political science and public administration journals on the use of lotteries to select beneficiaries of government programs consisted of a proposal to establish a draft for a mandatory national service to enhance social cohesion. (Galston 2012). One recent paper in political science evaluated the effects of candidate selection lotteries in Mexico on voter participation (Poertner, 2023). Neither directly assessed perceptions about actual lotteries.

Galston, W.A. (2012), "Citizenship and Civic Attachment: The Case for a Universal Service Lottery." Public Administration Review, 72:322-323. doi.org/10.1111/j.1540-6210.2012.02577.x

Poertner, Mathias. (2023). "Does Political Representation Increase Participation? Evidence from Party Candidate Lotteries in Mexico." American Political Science Review 117(2): 537–56. doi.org/10.1017/S000305

Appendix B Housing Lotteries Rules

In the earlier two of the four calls examined in the empirical sections of this paper, both held in 2011, applicants whose last two digits in their sequential number corresponded to the last two digits of the 1^{st} and 2^{nd} prize drawings were considered winners.

In the third call, applicants whose last three digits of the sequential number corresponded to the last three digits of the 1^{st} to 4^{th} prizes were pre-selected, while in the fourth call the winners were defined as those whose last three digits matched those of the 1^{st} prize or those whose last four digits matched those of the 2^{nd}, 3^{rd}, or 4^{th} prizes.

Appendix C Lotteries Survey Sampling

At the time in which this project was planned more than 50 general public MCMV lotteries had been held in the city of Rio de Janeiro since 2011. Lotteries 17/2016 and 20/2016 were used as the population from which the recent lotteries survey was selected and Lotteries 03/2011 and 06/2011 were used as the population from which the early lotteries survey was selected. Figure C1 shows a timeline for the lotteries and the surveys, and Table C1 provides additional information on the lotteries used in each survey. Very few individuals in one of recent lotteries were placed in a housing project whose units began to be delivered during our survey. We inadvertently interviewed four of these individuals but removed them from our data prior to the analysis.

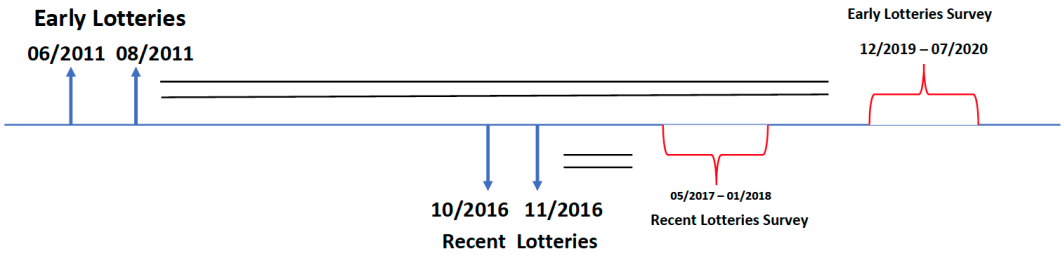


Figure C1: Timeline of Lotteries and Surveys

Table C1: MCMV Lotteries Included in the Quasi-Experimental Analyses

Survey	Early Lotteries Survey		Recent Lotteries Survey	
Field Dates	Dec. 2019 – July 2020		May 2017 – Jan. 2018	
Call	03-2011	06-2011	17-2016	20-2016
Lottery Date	06/11/2011	08/13/2011	10/19/2016	11/05/2016
Housing Projects	Park Imperial, Park Royal, Destri, Toledo, Rio Bonito, Estoril	Sevilla, Taroni, Toledo, Cascais	Vila Carioca, Safira, Sabóia	Porto Fino, Ametista
Non-Winners	295,235	318,789	580,983	484,151
Winners	2,983	6,505	2,299	612
Delivery Dates	2011/12 – 2012/11		2017/12 – 2018/12	
Pre-Sample	22,157		8,032	
Contacted	3,772		1,283	
Interviewed	2,119		795	

Notes: "Contacted" is defined as whether the selected person was found by our field team.

As C1 shows, each survey included two separate lotteries taken a few weeks apart. Most individuals that participated in one, participated in the other. Moreover, the number of units being assigned in each lottery varied, and so do the number of participants. Hence, probability of selection varied across lotteries too. In order to account for these irregularities we include sampling weights in our analyses of the recent lotteries survey to account for endogeneity induced by our sampling process.

Treatment assignment takes place at the lottery level, which means that, conditional on applying for lottery j, winning a lottery j is randomly determined. However, in our recent-lotteries survey (which was the first of the two surveys that we carried out), instead of sampling at the lottery-level, we selected respondents based on ever winning (not ever winning) any lottery of the two lotteries from which we were sampling rather sampling on winning (not winning) a given lottery j. This could lead to endogeneity because ever winning (not winning) depends on a series of unobserved factors (date in which subjects signed up, bureaucractic selection of individuals included in lotteries roll, etc.) that could affect being included in a lottery roll. Since we sample all winners, all winners get a weight of 1 and all non-winners get a weight of 1/probability of sampling a non-winner.

Appendix D Summary Statistics – Quasi-Experimental Analysis

The outcomes listed in Table 1 were constructed from survey items. Tables D1 and D2 report summary statistics for all items measured through our surveys that are referred to in the paper. Wording for all questions is reported in Appendix E. In the analyses reported in the paper, we standardized all outcomes as binary outcomes in order to facilitate presentation and interpretation. In Tables D1 and D2, we also show descriptive statistics in the original scale (below the dashed line).

Table D1: Descriptive Statistics – Recent Lotteries

	N	Min.	Mean	Median	Std. Dev.	Max	Miss.
Treatment assignment	1267	0.00	0.23	0.00	0.42	1.00	0
Compliance	1267	0.00	0.18	0.00	0.39	1.00	0
Exp. Compliance	1267	0.00	0.21	0.00	0.41	1.00	0
Sex	1267	0.00	0.30	0.00	0.46	1.00	0
Race (binary)	1267	0.00	0.25	0.00	0.44	1.00	0
Religion (binary)	1267	0.00	0.85	1.00	0.36	1.00	0
Children	1267	0.00	0.90	1.00	1.24	11.00	0
Schooling (years)	1267	0.00	9.01	11.00	4.14	15.00	0
National registry	1267	0.00	0.16	0.00	0.37	1.00	0
Formal job	1267	0.00	2.61	3.00	2.20	5.00	0
Log. avg. inc. from formal job	1267	0.00	4.35	6.28	3.23	9.13	0
Approves of lotteries (binary)	1160	0.00	0.49	0.00	0.50	1.00	107
Approves Lottery (alt. bin.)	1160	0.00	0.58	1.00	0.49	1.00	107
Improves lives (binary)	1192	0.00	0.84	1.00	0.36	1.00	75
Hope better lives (binary)	1257	0.00	0.95	1.00	0.22	1.00	10
Homeownership Dream (binary)	1263	0.00	0.95	1.00	0.21	1.00	4
Approves lottery	1160	-2.00	0.10	0.00	1.63	2.00	107
Improves lives	1192	-2.00	1.34	2.00	0.98	2.00	75
Homeownership dream	1263	-2.00	1.67	2.00	0.68	2.00	4

Notes: Age was measured using data from private vendors, employment, and earnings from employment was measured using RAIS data, and Registry was measured using CadUnico data. Treatment assignment was obtained by the researchers by digitizing lottery rolls and results from public sources.

Table D2: Descriptive Statistics - Early Lotteries

	N	Min.	Mean	Median	Std. Dev.	Max	Miss.
Treatment Assignment	3923	0.00	0.23	0.00	0.42	1.00	0
Compliance	3923	0.00	0.12	0.00	0.32	1.00	0
Age	3923	27.59	49.46	47.64	11.33	87.85	0
Sex	3923	0.00	0.38	0.00	0.49	1.00	0
Race (binary)	3923	0.00	0.24	0.00	0.43	1.00	0
National registry	3923	0.00	0.16	0.00	0.37	1.00	0
Formal job	3923	0.00	4.43	5.00	3.17	8.00	0
Log. avg. inc. from formal job	3923	0.00	4.68	5.92	2.67	8.67	0
Approves lotteries (binary)	3547	0.00	0.40	0.00	0.49	1.00	376
Approves lotteries (alt. bin.)	3547	0.00	0.46	0.00	0.50	1.00	376
Improves lives (binary)	3480	0.00	0.83	1.00	0.37	1.00	443
Hope for better life (binary)	3923	0.00	0.95	1.00	0.21	1.00	0
Homeonwership dream (binary)	3883	0.00	0.91	1.00	0.29	1.00	40
Approve lotteries	3547	-2.00	-0.27	-1.00	1.58	2.00	376
Improves lives	3480	-2.00	1.13	2.00	1.21	2.00	443
Would move	3923	0.00	0.58	1.00	0.49	1.00	0
Homeownership dream	3883	-2.00	1.37	2.00	0.89	2.00	40

Notes: Age was measured using data from private vendors, employment, and earnings from employment was measured using RAIS data, and Registry was measured using CadUnico data. Treatment assignment was obtained by the researchers by digitizing lottery rolls and results from public sources.

Appendix E Original Survey Question Wording

We reported translated versions of the survey questions in Table 1. Here we report the original wording, in Portuguese, of the survey questions used to construct the outcome variables in the paper:

Lotteries Fairness: Você diria que o sorteio é uma forma justa para selecionar participantes do programa Minha Casa Minha Vida? Muito justa, Pouco justa, Nem justa nem injusta, Pouco injusta, Muito injusta, NS, NR (recusa)

Improved Lives: Você acha que o Minha Casa Minha Vida piorou ou melhorou a qualidade de vida das pessoas que participam do programa? Melhorou muito, Melhorou um pouco, Não fez diferença, Piorou um pouco, Piorou muito, NS/NR (recusa)

Hope for better life Você diria que o programa Minha Casa Minha Vida traz esperança para as famílias de uma vida melhor? Traz esperança, Não faz diferença, Não, muito pouca gente ganha, Não, as moradias e localização das moradias são muito ruins, NS, NR (recusa) [In the early lotteries questionnaire the choices were Sim and Não]. In both waves, the answer was spontaneous (choices were not read aloud by interviewer)

Homeownership is a Dream: Tem gente que diz que ter a casa própria é um sonho, algo que transforma a vida das pessoas. Você concorda com isso? Concordo completamente, Concordo, Nem concordo nem discordo, Discordo, Discordo completamente, NS, NR (recusa)

Appendix F Balance Tests – Quasi-Experimental Analysis

We conducted a series of balance tests to assess the assumption that winners are similar to non-winners regarding pre-treatment covariates. As reported in Appendix F, these tests did not indicate meaningful systematic pre-treatment differences between the two groups. These results are consistent with studies using administrative data that also found no differences between winners and nonwinners on pre-treatment characteristics (Chagas and Rocha 2019, Souza 2019).

We present balance tests for the joint null hypothesis (by regressing the treatment assignment indicator on the pre-treatment covariates) as well as the test for each covariate. For the balance tests regarding the recent lotteries, we also include the same survey weights we included in the main analysis. Also, for pooled analyses (including both lotteries in each survey), we either include lottery fixed effects or follow Lin 2013.

We present balance tests for the joint null hypothesis (by regressing the treatment assignment indicator on the pre-treatment covariates) as well as the test for each covariate. For the balance tests regarding the recent lotteries, we also include the same survey weights we included in the main analysis. Also, for pooled analyses (including both lotteries in each survey), we either include lottery fixed effects or follow Lin (2013).

Table F1: Wald Test for Joint Null Hypothesis Test for Balance: Pooled Recent Lotteries

Res.Df	Df	F	Pr(>F)
1,256			
1,264	-8	1.54	0.1380

Note: Regression of treatment assignment on pre-treatment covariates (null model includes age). All standard errors are clustered at the individual level. Permutation p-value = 0.15

Table F2: F-Test for Joint Null Hypothesis Test for Balance: Pooled Early Lotteries

F	Df	p-value
0.49	6.00	0.82

Note: Regression of treatment assignment on pre-treatment covariates. All standard errors are clustered at the individual level. Permutation p-value = 0.799

Table F3: Pooled Balance Tests Recent Lotteries Survey: Covariates

	Estimate	Std. Error	t value	Pr(>|t|)	N
Sex (male)	0.05	0.04	1.19	0.24	1267
Race (white)	-0.01	0.04	-0.33	0.74	1267
Religion (any)	0.04	0.03	1.24	0.21	1267
Children (N)	-0.06	0.10	-0.62	0.53	1267
Schooling (years)	0.61	0.33	1.88	0.06	1267
Registry (CadUnico)	-0.04	0.03	-1.32	0.19	1267
Yrs in Formal Employment	0.39	0.17	2.26	0.02	1267
(Logged) Av. Formal Wages	0.50	0.23	2.12	0.03	1267
Not conditioning on age					
Sex (male)	0.04	0.04	0.95	0.34	1267
Race (white)	-0.01	0.04	-0.29	0.77	1267
Religion (any)	0.04	0.03	1.27	0.20	1267
Children (N)	-0.10	0.10	-0.98	0.33	1267
Schooling (years)	0.34	0.36	0.94	0.35	1267
Registry (CadUnico)	-0.04	0.03	-1.38	0.17	1267
Yrs in Formal Employment	0.27	0.19	1.41	0.16	1267
(Logged) Av. Formal Wages	0.29	0.27	1.07	0.28	1267

Notes: All standard errors are clustered at the individual level.

Table F4: Pooled Balance Tests Early Lotteries Survey: Covariates

	Estimate	Std. Error	t value	Pr(>|t|)	N
Lin Model					
Age	0.42	0.41	1.02	0.31	3923
Sex (male)	-0.02	0.02	-0.95	0.34	3923
Race (white)	0.00	0.01	0.15	0.88	3923
Registry (Cadunico)	-0.00	0.01	-0.34	0.73	3923
Years in Formal Employment	0.06	0.11	0.51	0.61	3923
Avg. Formal Wages	0.02	0.09	0.18	0.85	3923
Fixed Effects Model					
Age	0.33	0.40	0.85	0.40	3923
Sex (male)	-0.01	0.02	-0.43	0.66	3923
Race (white)	0.00	0.01	0.12	0.90	3923
Registry (Cadunico)	-0.01	0.01	-0.70	0.48	3923
Years in Formal Employment	0.10	0.11	0.96	0.34	3923
Avg. Formal Wages	0.07	0.09	0.72	0.47	3923

Notes: All standard errors are clustered at the individual level.

Table F5: Wald Test for Joint Null Hypothesis Test for Balance: Individual Lotteries – Recent Lotteries Survey

Lottery 17/2016				Lottery 20/2016			
Res.Df	Df	F	Pr(>F)	Res.Df	Df	F	Pr(>F)
786				461			
794	-8	1.52	0.1464	469	-8	0.49	0.8642

Notes: Regression of treatment assignment on pre-treatment covariates (null model includes age). All standard errors are clustered at the individual level. Permutation p-value 0.17 (Lottery 17/2016) and permutation p-value 0.89 (Lottery 20/2016)

Table F6: Wald Test for Joint Null Hypothesis Test for Balance: Individual Lotteries - Early Lotteries Survey

Lottery 03/2011				Lottery 06/2011			
Res.Df	Df	F	Pr(>F)	Res.Df	Df	F	Pr(>F)
1846				2063			
1852	-6	0.76	0.5979	2069	-6	0.74	0.6172

Notes: Regression of treatment assignment on pre-treatment covariates. All standard errors are clustered at the individual level. Permutation p-value 0.633 (Lottery 03/2011) and Permutation p-value 0.604 (Lottery 06/2011)

Table F7: Lottery 17/2016 Balance Tests: Covariates

	Estimate	Std. Error	t value	Pr(>|t|)	N
Sex (male)	0.07	0.04	1.77	0.08	796
Race (white)	-0.03	0.03	-0.94	0.35	796
Religion (any)	0.04	0.03	1.58	0.11	796
Children (N)	-0.02	0.10	-0.19	0.85	796
Schooling (years)	0.66	0.28	2.36	0.02	796
Registry (CadUnico)	-0.06	0.03	-2.33	0.02	796
Yrs in Formal Employment	0.24	0.17	1.46	0.14	796
(Logged) Av. Formal Wages	0.34	0.24	1.45	0.15	796
No conditioning on age					
Sex (male)	0.05	0.04	1.41	0.16	796
Race (white)	-0.03	0.03	-0.93	0.35	796
Religion (any)	0.05	0.03	1.73	0.08	796
Children (N)	-0.06	0.09	-0.63	0.53	796
Schooling (years)	0.25	0.33	0.78	0.44	796
Registry (CadUnico)	-0.06	0.03	-2.27	0.02	796
Yrs in Formal Employment	0.07	0.17	0.39	0.70	796
(Logged) Av. Formal Wages	0.06	0.26	0.24	0.81	796

Notes: All standard errors are clustered at the individual level.

Table F8: Lottery 20/2016 Balance Tests: Covariates

	Estimate	Std. Error	t value	Pr(>|t|)	N
Sex (male)	0.02	0.07	0.34	0.73	471
Race (white)	0.01	0.07	0.16	0.87	471
Religion (any)	0.03	0.05	0.60	0.55	471
Children (N)	-0.11	0.16	-0.68	0.50	471
Schooling (years)	0.55	0.57	0.96	0.34	471
Registry (CadUnico)	-0.01	0.05	-0.25	0.81	471
Yrs in Formal Employment	0.58	0.30	1.95	0.05	471
(Logged) Av. Formal Wages	0.69	0.38	1.81	0.07	471
No conditioning on age					
Sex (male)	0.02	0.07	0.29	0.77	471
Race (white)	0.01	0.07	0.20	0.84	471
Religion (any)	0.03	0.05	0.54	0.59	471
Children (N)	-0.14	0.16	-0.87	0.38	471
Schooling (years)	0.44	0.63	0.70	0.48	471
Registry (CadUnico)	-0.02	0.05	-0.38	0.70	471
Yrs in Formal Employment	0.51	0.33	1.55	0.12	471
(Logged) Av. Formal Wages	0.56	0.45	1.24	0.21	471

Notes: All standard errors are clustered at the individual level.

Table F9: Early Lotteries Balance Tests: Covariates

	Estimate	Std. Error	t value	Pr(>|t|)	N
Lottery 03/2011					
Age	0.79	0.75	1.05	0.30	1853
Sex (male)	-0.05	0.03	-1.80	0.07	1853
Race (white)	0.00	0.03	0.15	0.88	1853
Registry (Cadunico)	0.01	0.02	0.60	0.55	1853
Years in Formal Employment	-0.16	0.21	-0.75	0.45	1853
Avg. Formal Wages	-0.20	0.18	-1.11	0.27	1853
Lottery 06/2011					
Age	0.09	0.54	0.16	0.87	2070
Sex (male)	0.02	0.02	0.82	0.42	2070
Race (white)	0.00	0.02	0.02	0.99	2070
Registry (Cadunico)	-0.02	0.02	-1.29	0.20	2070
Years in Formal Employment	0.25	0.15	1.63	0.10	2070
Avg. Formal Wages	0.21	0.13	1.66	0.10	2070

Notes: All standard errors are clustered at the individual level.

Appendix G Attrition – Quasi-Experimental Analysis

Another important concern for the analysis of survey data is attrition. As shown in Appendix C, we had high rates of non-response because we were unable to find many respondents (as was City Hall). We are confident that we took exhaustive measures to contact respondents. We purchased contact (phone) information from three vendors; we attempted to contact each respondent at least three times on as many phone numbers as were associated with their identification; and in several instances, we located respondents who had won the lottery but City Hall had failed to contact. We also sought to minimize the possibility that we could induce different response rates between winners and non-winners by making our enumerators blind to treatment and by randomly determining the order of contact, and we approached winners and nonwinners using the same protocol. Moreover, we did not find differential attrition rates between winners and non-winners (see Appendix G).

In addition to the comparison of response rates between winners and non-winners, we also compared our survey sample estimates to administrative data benchmark in which there is no attrition. We compared the effect of winning MCMV on having a formal job and earnings from a formal job in our survey sample and in the administrative data for these same lotteries (623,598 individual-

lottery observations). Results of winning MCMV on having a formal job and earning from that job using our survey data, with attrition, are statistically indistinguishable from the results found using administrative data, without attrition (Appendix G). While there is no way of definitely knowing whether missingness is correlated to potential outcomes, these results are reassuring.

We compare attrition patterns by regressing indicators for response and response conditional on "picking up" (the phone call) indicator on treatment assignment, respondent's sex (which we estimated based on subject's name using R package `genderBR`), pre-treatment earnings from a formal job, having a formal job, and interactions between treatment and these covariates. We then estimate a F-test of the hypothesis that all interaction coefficients are zero. In all models, we cluster standard error at the respondent level and we include sampling weights for the analysis of the 2011 survey. We find no systematic difference between in attrition patterns between treatment and control groups in both surveys.

In Tables G3 and G4 we show comparisons of attrition rates across treatment arms. In Columns (1) and (2) of both Tables, we regress a response indicator on treatment assignment, using Lin's (2013) model and fixed-effects for each lottery, respectively. In columns (3) and (4) of both Tables, we compare attrition rates across treatment and control by regression a response indicator conditional of picking up the phone on treatment assignment. In all models, we cluster standard errors at the respondent level and we include sampling weights for the analysis of the 2011 survey. Overall, we do not observe different response rates across treatment and control groups even though there is some evidence of in models (1) and (2) of Table G3, at the 10% level of statistical significance. Unfortunately, we do not have age information available for all subjects we attempted (only for those interviewed). This means that we cannot examine attrition conditional on age, as we did for our main analyses in the paper, since randomization was (marginally) conditional on age in the lotteries conducted in 2016.

Table G1: Wald Test for Attrition Patterns: Survey Recent Lotteries

Interviewed Indicator				Interviewed Conditional on Picking up			
Res.Df	Df	Chisq	Pr(>Chisq)	Res.Df	Df	Chisq	Pr(>Chisq)
11,716				1,533			
11,719	-3	2.53	0.4701	1,536	-3	0.68	0.8770

Notes: All standard errors are clustered at the individual level.

Table G2: Wald Test for Attrition Patterns Survey Early Lotteries

Interviewed Indicator				Interviewed Conditional on Picking up			
Res.Df	Df	Chisq	Pr(>Chisq)	Res.Df	Df	Chisq	Pr(>Chisq)
29,624				6,432			
29,627	-3	2.39	0.4950	6,435	-3	0.18	0.9808

Notes: All standard errors are clustered at the individual level.

Table G3: Differential Attrition Rates: Survey with Recent Lotteries

	Resp. Lin Model	Resp. Fixed Effects Model	Resp. | Picking up Lin Model	Resp. | Picking up Fixed Effects Model
(Intercept)	0.10***		0.66***	
	(0.00)		(0.02)	
Winning Lottery	−0.02	−0.02	−0.12	−0.17
	(0.01)	(0.01)	(0.08)	(0.13)
Lottery ID	−0.00		−0.02**	
	(0.00)		(0.01)	
Winning Lottery ×	−0.02		0.17	
Lottery ID	(0.02)		(0.16)	
R^2	0.00	0.00	0.00	0.00
Adj. R^2	−0.00	−0.00	−0.00	−0.00
Num. obs.	12,594	12,594	1,772	1,772
RMSE	3.06	3.06	4.81	4.81
N Clusters	7,764	7,764	1,096	1,096

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table G4: Differential Attrition Rates: Survey with Early Lotteries

	Resp. Lin Model	Resp. Fixed-Effects Model	Resp. | Picking up Lin Model	Resp. | Picking up Fixed-Effects Model
(Intercept)	0.12***		0.56***	
	(0.00)		(0.01)	
Winning Lottery	−0.00	−0.00	0.00	0.01
	(0.00)	(0.00)	(0.01)	(0.01)
Lottery Id	0.00		−0.01	
	(0.00)		(0.01)	
Winning Lottery ×	0.01		0.05	
Lottery ID	(0.01)		(0.03)	
R^2	0.00	0.00	0.00	0.00
Adj. R^2	−0.00	−0.00	0.00	−0.00
Num. obs.	31,732	31,732	6,947	6,947
RMSE	0.33	0.33	0.50	0.50
N Clusters	17,117	17,117	3,771	3,771

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

In an effort to assess whether nonresponse could be leading to biased estimates, we compare estimates for administrative outcomes obtained using our survey sample against a "benchmark" estimates for the same outcomes using the whole population for which these data are observed and that does not suffer from nonresponse (attrition). Specifically, we estimate the effect of winning a lottery in 2011 on earnings in our survey sample and in the whole population. Table G5 shows that, in both the survey sample and in the whole population there is a null effect of winning a lottery on these labor outcomes. Furthermore, estimates are statistically indistinguishable from each other (p-value = 0.66, earnings outcomes; p-value = 0.40, formal job outcome). Unfortunately we do not have any post-treatment outcome using administrative data for the survey with 2016 lottery participants.

Table G5: Comparison of Survey Estimates to "Benchmark" Administrative Data Estimates

	Administrative Average (logged) Earnings	Administrative Formal Job	Survey Average (logged) Earnings	Survey Formal Job
Winning a lottery	−0.03	−0.01	0.03	0.02
	(0.04)	(0.01)	(0.12)	(0.03)
R^2	0.00	0.00	0.00	0.00
Adj. R^2	0.00	0.00	−0.00	−0.00
Num. obs.	623,410	623,410	3,923	3,923
RMSE	3.53	0.95	3.47	0.93
N Clusters	350,689	350,689	2,119	2,119

Notes: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Appendix H Outcome Variable Descriptives – Quasi-Experimental Analysis

In the main body of the paper we mentioned correlation coefficients between the outcomes and the overall distribution of outcomes. In this section, we present the complete descriptives, omitted from the main body of the paper for parsimony.

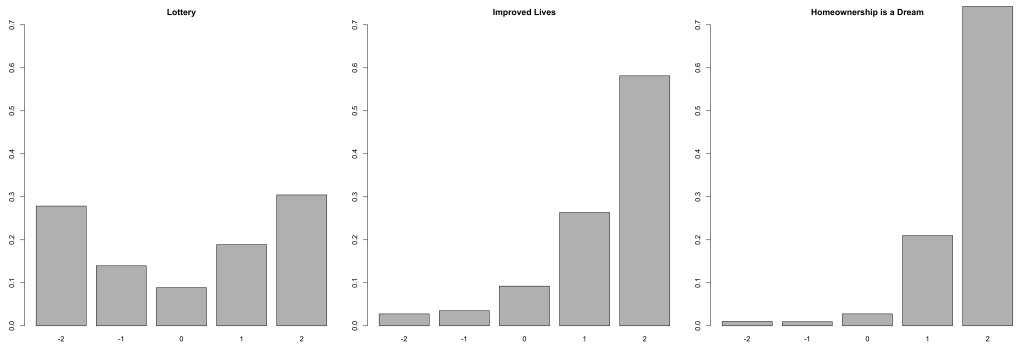
Appendix H.1 Correlation between outcomes

The outcomes described in Table 1 capture different aspects of evaluation of the MCMV. While the prospective and retrospective evaluations of the MCMV are moderately associated to each other (r=0.32 and r=0.26, in each wave), the association between the other outcomes is never particularly strong. Evaluations of the lottery, in particular, are positively but weakly associated with the other three outcomes in both waves (r=[0.27, 0.14, 0.11] and r=[0.22, 0.08, 0.05]). In sum, there is suggestive evidence that individuals' evaluations of the lotteries is reflective of processes and attitudes different the ones driving the evaluation than other outcomes we examine here.

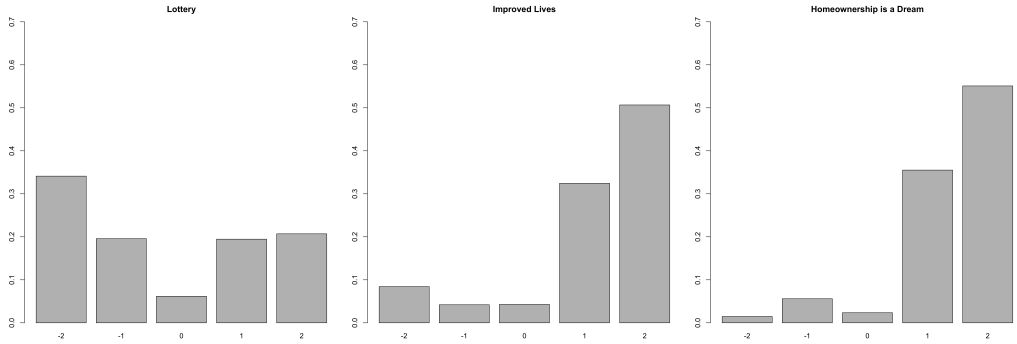
Appendix H.2 Polarization

In Figure H1 we present the data in support of this claim by showing the outcomes in their original scale. We do not report results for the prospective evaluation question because it was originally coded in as binary variable. The figure shows that approval of the lottery has a clear

bimodal distribution of answers, with fewer than 10% of respondents picking the middle/neutral category.



(a) Recent Lotteries (Lotteries) (b) Recent Lotteries (Lives) (c) Recent Lotteries (Dream)



(d) Early Lotteries (Lotteries) (e) Early Lotteries (lives) (f) Early Lotteries (Dream)

Figure H1: Distribution of Raw Outcomes

Notes: Distributions reflect the outcome variables original scales. Hope for better life is omitted because it was originally coded on a two-point scale. See Table 1 for details.

Appendix I Tables of Results — Quasi-Experimental Analysis

Figure 1, in the main body of the paper, depicts the ITT estimates of the effect of winning the lottery on each outcome of interest. In this section of the appendix we provide the numerical estimates.

Table I1: Table for Figure 1a, Wave 1 Results for ITT
with fixed effects and without controls (except age).

	Goal	Retro	Prosp.	Just
Lottery Winner	−0.022	0.051+	0.019	0.208***
	(0.017)	(0.027)	(0.016)	(0.038)
Age	−0.001	−0.003*	−0.002*	0.000
	(0.001)	(0.001)	(0.001)	(0.002)
Num.Obs.	1263	1192	1257	1160
R2	0.000	0.017	0.011	0.027
AIC	761.0	2285.4	1162.4	2883.3
BIC	776.5	2300.6	1177.9	2898.4
RMSE	0.21	0.36	0.22	0.49
Std.Errors	by: fieldID	by: fieldID	by: fieldID	by: fieldID

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Survey weights included. Clustered standard errors.

Table I2: Table for Figure 1a, Wave 1 Results for ITT with fixed effects and with controls.

	Goal	Retro	Prosp.	Just
Lottery Winner	−0.020	0.057*	0.027	0.218***
	(0.018)	(0.027)	(0.018)	(0.038)
Age	−0.001	−0.006***	−0.003**	−0.003
	(0.001)	(0.002)	(0.001)	(0.002)
Sex (male)	−0.001	−0.018	−0.041	0.063
	(0.020)	(0.039)	(0.026)	(0.054)
Race (white)	0.035*	0.061	0.025	0.085
	(0.015)	(0.038)	(0.022)	(0.055)
Religion (any)	0.003	0.112*	−0.024	0.078
	(0.024)	(0.055)	(0.023)	(0.061)
Children (N)	0.004	0.009	−0.010	−0.009
	(0.005)	(0.011)	(0.013)	(0.018)
Schooling (years)	0.002	−0.013**	−0.005+	−0.022***
	(0.002)	(0.005)	(0.003)	(0.006)
Years in Formal Employment	−0.006	−0.004	−0.002	0.032
	(0.008)	(0.020)	(0.010)	(0.027)
(Logged) Av. Formal Wages	−0.001	0.000	−0.002	−0.023
	(0.005)	(0.014)	(0.007)	(0.019)
Num.Obs.	1263	1192	1257	1160
R2	0.002	0.036	0.024	0.057
AIC	760.7	2248.3	1147.3	2847.1
BIC	817.2	2304.2	1203.8	2902.7
RMSE	0.21	0.36	0.22	0.49
Std.Errors	by: resp.	by: resp.	by: resp.	by: resp.

Notes: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Survey weights included. Clustered standard errors.

Table I3: Table for Figure 1b, Wave 2 Results
for ITT with fixed effects and without controls.

	Goal	Retro	Prosp.	Fair
Lottery Winner	0.010	−0.001	−0.003	0.052**
	(0.010)	(0.013)	(0.008)	(0.018)
Num.Obs.	3883	3480	3923	3547
R2	0.000	0.000	0.000	0.002
R2 Adj.	0.000	−0.001	0.000	0.001
AIC	1458.1	3048.0	−974.7	5005.3
BIC	1470.6	3060.3	−962.1	5017.7
RMSE	0.29	0.37	0.21	0.49
Std.Errors	by: resp.	by: resp.	by: resp.	by: resp.

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Survey weights included. Clustered standard errors.

Table I4: Table for Figure 1b, Wave 2 Results for ITT with fixed effects and with controls.

	Goal	Retro	Prosp.	Fair
Lottery Winner	0.009	0.000	−0.003	0.054**
	(0.010)	(0.013)	(0.008)	(0.018)
Age	0.000	−0.005***	0.000	−0.001
	(0.001)	(0.001)	(0.000)	(0.001)
Sex (male)	−0.029*	−0.022	−0.011	0.101***
	(0.014)	(0.019)	(0.010)	(0.025)
Race (white)	−0.017	−0.022	−0.012	0.019
	(0.016)	(0.021)	(0.012)	(0.027)
Years in Formal Employment	0.005	0.010+	0.004	0.003
	(0.005)	(0.006)	(0.003)	(0.008)
(Logged) Av. Formal Wages	−0.004	−0.016*	−0.004	−0.002
	(0.006)	(0.008)	(0.004)	(0.010)
Num.Obs.	3883	3480	3923	3547
R2	0.005	0.020	0.003	0.013
R2 Adj.	0.003	0.017	0.001	0.011
AIC	1450.1	2991.1	−974.4	4977.7
BIC	1500.2	3040.4	−924.2	5027.1
RMSE	0.29	0.37	0.21	0.49
Std.Errors	by: resp.	by: resp.	by: resp.	by: resp.

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Survey weights included. Clustered standard errors.

Appendix J Additional Quasi-Experimental Analyses

In the main body of the paper we present ITT estimates obtained with the lottery fixed-effects setup described in Equation 1. In this section we show that results obtained following the interactive specification proposed by Lin (2013) are all but identical. Lin's specification is defined as

follows:

$$Y_i = \beta_0 + \beta_1 Z_{ij} + \sum_{m=1}^{M} (\mu_m \mathbf{X}_i + \sigma_m Z_{ij} \times \mathbf{X}_i) + \sum_{j=1}^{J-1} (\Lambda_j B_{ji} + \gamma_j Z_{ij} \times B_{ji}) + u_{ij} \tag{2}$$

where Y_i is an outcome measure for an applicant i, $Z_{ij} = 1$ represents a winning applicant i in lottery j and $Z_{ij} = 0$ represents a non-winning applicant i in lottery j. The coefficient β_1 represents the effect of winning a lottery.

Table J1: Wave 1 Results for ITT via Lin (2013) without controls, panel (a), and with controls, panel (b).

	Estimate	Std. Error	p-value	N	Std. Estimate
Approves lotteries	0.22	0.04	0.00	1160	0.45
Improves lives	0.05	0.03	0.06	1192	0.15
Hope better life	0.02	0.02	0.26	1257	0.09
Homeownership dream	-0.04	0.02	0.10	1263	-0.18
Approves lotteries	0.23	0.04	0.00	1160	0.46
Improves lives	0.06	0.03	0.05	1192	0.16
Hope better life	0.02	0.02	0.23	1257	0.10
Homeownership dream	-0.04	0.02	0.10	1263	-0.18

Table J2: Wave 2 Results for ITT via Lin (2013) without controls, panel (a), and with controls, panel (b).

	Estimate	Std. Error	p-value	N	Std. Estimate
Approves lotteries	0.06	0.02	0.00	3547	0.13
Improves lives	0.00	0.01	0.85	3480	0.01
Hope better life	-0.00	0.01	0.85	3923	-0.01
Homeownership dream	0.01	0.01	0.25	3883	0.04
Approves lotteries	0.06	0.02	0.00	3547	0.13
Improves lives	0.00	0.01	0.82	3480	0.01
Hope better life	-0.00	0.01	0.87	3923	-0.01
Homeownership dream	0.01	0.01	0.28	3883	0.04

Table J4: Wave 1 Results for ITT using alternative codings of Justness outcome (without controls)

	Just (alt. bin)	Just (orig. scale)
Lottery Winner	0.203***	0.742***
	(0.036)	(0.121)
Age	0.000	−0.002
	(0.002)	(0.006)
Num.Obs.	1160	1160
R2	0.001	0.001
R2 Adj.	−0.002	−0.001
Std.Errors	by: fieldID	by: fieldID

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table J3: Wave 2 Results for ITT using IPW, without controls

	Goal	Retro	Prosp.	Fair
Lottery Winner	0.012	0.003	−0.001	0.062***
	(0.010)	(0.014)	(0.008)	(0.018)
Num.Obs.	3883	3480	3923	3547
R2	0.000	0.000	0.000	0.002
Std.Errors	by: resp.	by: resp.	by: resp.	by: resp.

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix K Complier Average Treatment Effects – Quasi-Experimental Analysis

In all four lotteries, not all winners learn that they have won the lottery or become a beneficiary. Furthermore, some non-winners may believe that they have won a lottery or that they may have become beneficiaries via other lotteries or parts of the program. To account for these situations, in addition to intent-to-treat effect, we also estimate the effect of winning the lottery among compli-

Table J5: Wave 2 Results for ITT using alternative codings of Justness outcome (without controls)

	Just (alt. bin.)	Just (orig. scale)
Lottery Winner	0.045*	0.176**
	(0.018)	(0.057)
Num.Obs.	3547	3547
R2	0.001	0.002
R2 Adj.	0.001	0.002
Std.Errors	by: fieldID	by: fieldID

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

ers.

Compliance with treatment has a different definition in each survey. A the time of the recent lotteries survey, winners had not yet signed a MCMV agreement or received their new home. Hence, in this context, we defined compliance as "knowing about winning." While all lotteries are public and the results are published in the official gazette and on the internet, only lottery winners are notified directly via telegrams, and they are often not found by city officials (many winners were not found by our field team either). There is no direct notification to lottery non-winners. We measured "knowing about winning" by hand-coding whether the telegrams sent to lottery winners were received by them or a family member as well as by asking, in our survey, whether the subject had learned about winning an MCMV lottery. When we combine both measures, we find that 199 out of 286 individual-lottery winners learned about winning the lottery, and 69 out of 981 individual-lottery non-winners believed they were informed that they won the lottery. Non-winners sometimes participate in other social programs or are genuinely confused about government notifications, which leads some to declare having being notified as lottery winners. Importantly, our survey and telegram data are measured at the individual level. Therefore, compliance (D_i) is measured at the individual-level i and not individual-lottery ij level.

In the early lotteries survey, we define treatment as becoming a program beneficiary. We measure becoming a beneficiary based on whether the lottery participant or their spouse signed a MCMV contract with the CEF. We obtained information on all signers of MCMV agreements through an Access to Information request, and we used the national social assistance registry (*Cadunico*) to determine the spousal status of lottery winners. We found that 221 out of 912 individual-lottery winners became MCMV recipients and 249 out of 3,012 individual-lottery non-winners signed MCMV agreements. Individuals might have become MCMV beneficiaries through other lotteries than the ones we used in our survey or because of forced relocation by the city. Again, compliance (D_i) is measured at the individual i level.

In both surveys, we estimate complier average causal effects (CACE) through a two-stage least

squares procedure, employing the same pre-treatment controls and the fixed-effects framework defined in Equation 1.

Table K1: Table for Figure 1(a), Wave 1 Results for CACE
FE without controls, panel (a), and with controls, panel (b).

	Estimate	Std. Error	p-value	N	Std. Estimate
Approves lotteries	0.32	0.06	0.00	1160	0.65
Improves lives	0.08	0.04	0.06	1192	0.22
Hope better life	0.03	0.03	0.24	1257	0.13
Homeownership dream	-0.03	0.03	0.20	1263	-0.16
Approves lotteries	0.34	0.06	0.00	1160	0.69
Improves lives	0.09	0.04	0.04	1192	0.25
Hope better life	0.04	0.03	0.12	1257	0.19
Homeownership dream	-0.03	0.03	0.25	1263	-0.15

Table K2: Table for Figure 1(b), Wave 2 Results for CACE
FE without controls, panel (a), and with controls, panel (b).

	Estimate	Std. Error	p-value	N	Std. Estimate
Approves lotteries	0.29	0.10	0.00	3547	0.59
Improves lives	-0.01	0.08	0.94	3480	-0.01
Hope better life	-0.02	0.05	0.71	3923	-0.08
Homeownership dream	0.06	0.06	0.34	3883	0.20
Approves lotteries	0.29	0.10	0.00	3547	0.59
Improves lives	-0.01	0.08	0.94	3480	-0.01
Hope better life	-0.02	0.05	0.71	3923	-0.08
Homeownership dream	0.06	0.06	0.34	3883	0.20

Among "compliers," which in the recent surveys corresponds to participants who were aware of the result of the lottery, the effect is even larger (0.32). The effects for individuals who participated in MCMV lotteries several years earlier are similar in early and recent lotteries surveys (about 0.60 in standardized effect sizes).

Appendix L Qualitative Semi-structured Interviews

Our in-depth interviews with 15 participants of the recent lotteries survey were conducted in August and September, 2018. We selected a sample of individuals we had interviewed in our survey and invited them for a in-depth conversation, in the hopes of interviewing 10 winners and 5 non-winners. Our sample was composed of three groups, two ensure that we would have relevant variation in our set on interviewees: i) subjects who evaluated former president Lula higher than the sample average, ii) lower than the sample average (we selected these two former groups with higher probability) and iii) average. We used president Lula's evaluation because it is a key outcome in other parts of our project. We offered compensation for subjects' time and transportation to the place where we conducted interviews (we did not offer any compensation for the survey interviews).

In our analysis of both MCMV and TETO interviews (see below), we developed a set of five themes to more systematically classify each interview focusing on whether i) the subject understood the lottery system (and, if not, what was their understanding), ii) whether the subject expressed a consequentialist view in support of the lottery, iii) if lottery (or whatever method they believed was implemented) was just in principle, iv) whether it was fair as implemented, and v) what was a preferable alternative to lotteries (or whatever methods they believed were implemented). The goal of these themes was to create a set of common classifications across interviews. Three of these themes (iii, iv, v) were identified ex-ante based on our research questions and two themes (i, and ii) were identified inductively after our first iteration analyzing these interviews.

Our 50 in-depth TETO qualitative interviews were conducted in mid-2021. Subjects were invited and interviewed via phone. We interviewed a convenience sample (based on availability, owing a phone, and ability to receive our compensation via pre-paid phone) of subjects who agreed to be interviewed after having answered a survey questionnaire. Subjects were read a consent form prior to the interview and told that the interview would be recorded for research purposes. All subjects received a compensation for their time and they were paid via pre-paid phone credits.

We used two very similar interview scripts (one for winners and another one for non-winners). Scripts were used as a general guide (find the complete scripts, in Portuguese, in our replication materials). We explored answers to two main questions in the interview prompt, and to follow-up clarification questions that followed: "O que você acha de escolher os beneficiários do Minha Casa Minha Vida através de sorteio?" (What do you think of selecting MCMV beneficiaries through a drawing) and "Você concorda ou discorda com essa prática?" (Do you agree or disagree with this procedure?)

We also reviewed several questions related to applicants' views of the application process, which are broken down into multiple questions on how/when/why people signed up for MCMV.

For TETO participants, we asked: "E você sabe como é que a Teto selecionou as pessoas para ganhar a casa? Você sabe como é que ela fez para selecionar as pessoas?" (Do you know how TETO selected recipients for a new home?) and "Por último, eu queria saber sobre a escolha de quem vai receber a casa da Teto. Você acha que é justa a maneira com que a Teto escolhe as pessoas que vão receber a casa, em geral?"(Lastly, we would like to ask about the choice of recipients. Do you think it was fair, how TETO did it?). Finally, we had questions to probe perceptions of discretion: "Você acha que pessoas que têm mais influência, que são mais conhecidas

na comunidade, elas podem ter mais vantagens na hora de conseguir a casa, ou não? / Claro que, no caso, cada caso é um caso, você recebeu a casa e tudo. Mas, em geral, você acha que quem tem contatos com alguém importante da comunidade tem alguma vantagem para receber a casa, mesmo com o sorteio?" (Do you think people who have more influence and are better known in the community, are more likely to receive a home? [For winners] We know it's a case by case matter, you yourself received a home. But, in general, do you think people with important contacts and connections is more more likely to receive a home, even with a lottery?)

The second set of interviewees were participants in a different housing program, implemented by a NGO called TETO in five states in Brazil. TETO builds prefabricated houses in slums. Different from MCMV, TETO's houses do not lead to homeownership, as TETO does not provide a title, and the NGO usually works with even more vulnerable families than MCMV. TETO's prefabricated houses, while much simpler than MCMV units, are a substantial improvement compared to their program participants' current dwellings. Yet TETO also selects beneficiaries via lottery, using a randomized wait-list in which a set of pre-selected participants enter a lottery that randomizes the order in which they are offered a home.If a participant does not accept the house or issues arise rendering a participant ineligible – such as construction not being legally or practically feasible in the proposed location – the next participant receives an offer. Winners are those who receive an offer of a TETO house. Importantly, only participants meeting a certain threshold in a vulnerability index created by the NGO are included in the lottery. While both MCMV and TETO have criteria in place to determine eligibility for the lottery, participants of TETO's program tend to be in a much more vulnerable situation than the average MCMV beneficiary and, among those eligible, the odds of being offered a TETO house are substantially larger than those of receiving a MCMV.

As with MCMV interviews, subjects did appear to understand the basic principles of a lottery. When asked open-ended questions about the lotteries' fairness, the appraisal of lotteries, in principle, varies considerably. In the TETO interviews, support for lotteries was conditional on homogeneity of those entering the lottery, as in the MCMV interviews. TETO applicants were much more homogeneous than MCMV applicants and support for the lotteries was also (qualitatively) higher in this group. One interviewee explained that she thought the lottery was fair because those "who don't need" the program were not allowed to enter the lottery:

T14: There are a lot of people who need it . . . I was happy because I was selected, but my neighbor, who also needs it, was not. So I am happy and sad at the same time. Do you understand? [. . .] All of those selected really needed it. There were people who, let's say, had better conditions and they also wanted to participate, but they [Techo] did not include them [. . .]

The perception that lotteries avoided preferential treatment was shared by applicants of TETO program as well, suggesting that the fear of being treated unfairly due to personal (or political) connections to the agents administering the benefits is not exclusive to government programs. For example, subject T11 explains why lotteries are fair based on the perception that they avoid preferential treatment:

T11: [. . .] Because I think this [lotteries] is an easy way. There are too many people. Many families. They [TETO] come, they interview everyone. Then what happens? They found who needed the most and had a lottery. They had to do it like that, to avoid confusion, so that there isn't someone who is selected, such as "This person chose that person because they are friends. Due to friendship." The TETO folks have lots of friends, they created lots of friendships here. So,

they did not decide based on friendship, they chose based on who they selected [via lottery]. So, they are right. No mess, no fighting, no talking that one was currying favor ...

The comparison between MCMV and Techo reveals many similarities in subjects' overall views of lotteries. These similarities are partly surprising given the differences in the ways the lotteries were implemented and the differences in the probability of winning the lottery (much higher for TETO participants than MCMV). But they are reassuring in suggesting that subjects' views on lotteries are not completely program specific. The key difference in the interviewees from MCMV and TETO lies in their views of whether lottery participants were "deserving." Like MCMV, TETO also has eligibility criteria for those entering the lottery, but these appear to be better communicated and clearer to participants, which resonated with several respondents when they were asked whether the lotteries were fair. The TETO lotteries were perceived as selecting from a pool of similarly worthy applicants.

Appendix M Design of Survey Experiment & Survey Sampling

The survey experiment was designed as a head-on comparison of two alternatives to selection of beneficiaries by lottery that were frequently mentioned in our qualitative interviews. The general structure of the experiment was a simple three-condition design with four outcome-related questions:

Introductory vinette (all)		
Imagine um programa habitacional no qual o governo entrega casas novas a famílias de baixa renda que pagam, em troca, pequenas prestações ao longo de muitos anos. Apenas famílias que recebem menos de um salário mínimo por mês podem se inscrever para receber uma casa. Mas, ainda assim, há muito mais gente inscrita do que casas disponíveis.		
Cond. 1: Need	Cond. 2: Queing	Cond. 3: Lottery
Para escolher quem dentre aqueles inscritos e com renda menor a um salário mínimo mensal vão receber a casa, o governo decidiu priorizar os de menor renda, vivendo em locais de risco ou cujas famílias incluam indivíduos com problemas graves de saúde e idosos.	Para escolher quem dentre aqueles inscritos e com renda menor a um salário mínimo mensal vão receber a casa, o governo decidiu priorizar as famílias que estiverem inscritos há mais tempo.	Para escolher quem dentre aqueles inscritos e com renda menor a um salário mínimo mensal vão receber a casa, o governo decidiu fazer um sorteio público pela loteria federal para escolher os ganhadores.

Outcomes: 1. Você acha que um programa desse tipo melhora ou piora a vida das pessoas? 2. Você acha essa maneira de decidir quem vai receber as casas é justa? 3. Você acha que essa maneira de

Table M1: Summary statistics I: experimental survey data

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Prog. Eval.	3	3	0.8	0.4	0.0	1.0	1.0
Justness	3	4	0.7	0.5	0.0	1.0	1.0
Efficiency	3	4	0.6	0.5	0.0	1.0	1.0
Favoritism	3	5	0.2	0.4	0.0	0.0	1.0
Age	67	0	43.3	15.8	17.0	42.0	84.0

escolher os beneficiários garante que os escolhidos sejam os que mais necessitam? 4. Você acha que pessoas que conhecem políticos, cabos eleitorais ou funcionários das prefeituras terão mais chances do que outros de conseguir uma casa?

The survey experiment was part of a larger survey conducted by Quaest. The survey was conducted face-to-face, with household visits by enumerators using tablets. The sample of 2,000 respondents is selected in three stages. First, a random selection of municipalities determines the municipalities that will enter the sample, and the probability of selection is proportional to the municipalities' voting population over 16 years old. The second stage is another random selection of census areas within the selected municipalities. In the last and third stage, Quaest selects households and individuals using sample quotas to mimic the national electorate, taking into account sex, age, formal schooling, income, and socioeconomic status.

Table M2: Summary statistics II: experimental survey data

		N	%
Exp. Conditions	Need-only	682	34.1
	Queue	633	31.6
	Lottery	685	34.2
Class	00-01SM	455	22.8
	01-02MW	340	17.0
	02-03MW	436	21.8
	03-05MW	311	15.6
	05-10SMW	369	18.4
	10+SM	89	4.5
Religion	Catholic	976	48.8
	Evangelical	573	28.6
	Other	366	18.3
	Other Religions	85	4.2
Sex	Fem	1049	52.5
	Male	951	47.5
Region	CO	148	7.4
	N	147	7.3
	NE	540	27.0
	S	296	14.8
	SE	869	43.5

Appendix N Balance Tests – Experimental Data

Based on our PAP, we assess balance based on income (high-income), religion (evangelicals), socioeconomic status (high SES), age (below median), region (South and/Southeast), and race (white). The dataset provided by the survey company did *not* include the SES variable, as we had originally imagined, so we restricted the analysis to the other variables.

We performed two types of balance analysis, one for joint balance by regression treatment indicators on all covariates and by examining each covariate separately. We conduct two F-tests by regressing an indicator of assignment to queue vs. need on all covariates (Table N1) and then by regression an indicator of assignment to lottery vs. need on all covariates (Table N2). We also examined joint balance by conducting a multinomial regression of the treatment indicator with all three categories (need, queue, lottery) on all covariates and comparing against a regression of the treatment indicator on a constant. All three models suggest our randomization was succcessfull (Table N3).

Table N1: Joint Test of Balance: Queue vs. Need

	Res.Df	Df	F	Pr(>F)
1	1308			
2	1314	-6	0.57	0.7566

Table N2: Joint Test of Balance: Lottery vs. Need

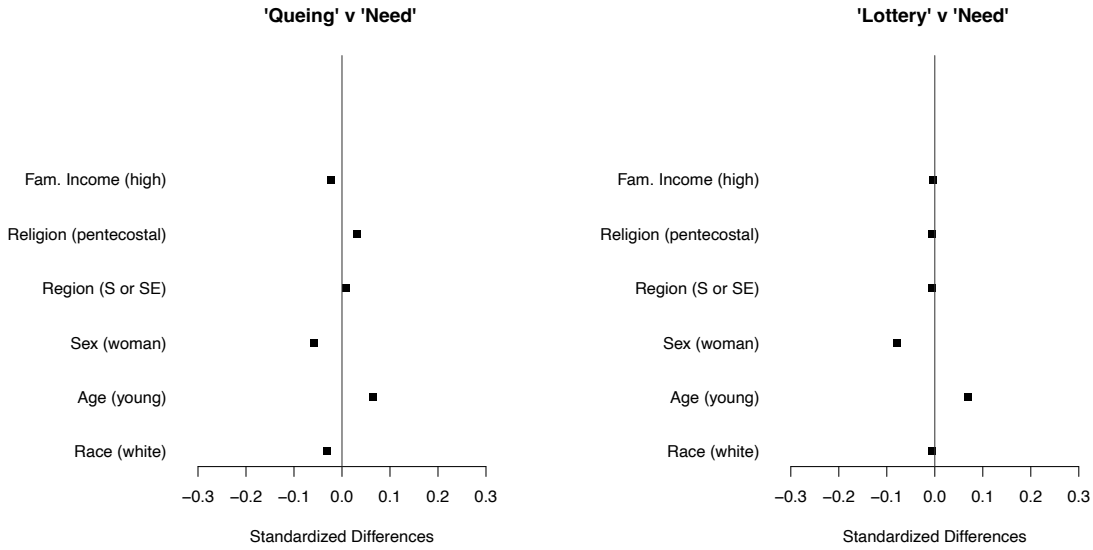
	Res.Df	Df	F	Pr(>F)
1	1360			
2	1366	-6	0.64	0.6998

Table N3: Joint Test of Balance: Multinomial model with treatment indicator (need, queue, lottery) as dependent variable

	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
1	1	3998.00	4391.87				
2	All Covariates	3988.00	4386.65	1 vs 2	10.00	5.22	0.88

Appendix O Additional Analyses – Experimental Data

In this section we reported additional specifications of the analyses of the experimental data, as well as a comparison of treatment effects across specifications.



(a) Baseline v. Cond. 2 (b) Baseline v. Cond. 3

Figure N1: Balance on Pre-treatment Observables

Appendix O.1 Heterogenous treatment effects

Table O1: Experimental results (all outcomes)

	Main Effects				Het. Effects	
	Prog. Eval	Justness	Efficiency	Favoritism	Just|Age	Just|Income
Need (baseline)	**0.90*****	**0.80*****	**0.70*****	**0.18*****	**0.83*****	**0.79*****
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Queue	**−0.04***	**−0.09*****	−0.04	−0.01	**−0.09****	**−0.08****
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Lottery	**−0.19*****	**−0.26*****	**−0.10*****	0.02	**−0.28*****	**−0.29*****
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Young					−0.05	
					(0.03)	
Young×Queue					−0.00	
					(0.05)	
Young×Lottery					0.04	
					(0.05)	
High. Inc						0.03
						(0.03)
High. Inc×Queue						−0.02
						(0.05)
High. Inc×Lottery						**0.09**$^{\cdot}$
						(0.05)
Num. obs.	1936	1924	1917	1908	1924	1924

Notes: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$. Robust standard errors. No pre-treatment covariates included.

Table O2 reports the results for the same analyses that were reported in Table 2, in the main body of the paper, but estimated with the pre-registered set of controls. We included categorical variables that controlled for region (five levels), religion (three levels), income (six levels), religion (three levels), sex (two levels), and a continuous variable for age. Coefficients for these control variables are not reported for parsimony.

As expected by the nature of these control variables and by random assignment, we see different

values in the baseline coefficients, but the substantive interpretation of the treatment effects for the different conditions is all but identical to what we reported in the main body. The program with selection by lottery has lower overall evaluation, the lottery as a selection mechanism is considered substantially less fair and somewhat less efficient. The lottery is marginally better in reducing favoritism, but the effect here is not statistically significant.

Table O2: Results with controls (all outcomes)

	Main Effects				Het. Effects	
	Prog.Eval	Fairness	Efficiency	Favoritism	Fair|Age	Fair|Income
Need-only (baseline)	**1.06*****	**0.72*****	**0.86*****	**0.09**$^{\cdot}$	**0.75*****	**0.78*****
	(0.04)	(0.06)	(0.06)	(0.05)	(0.09)	(0.05)
Queue	**−0.04***	**−0.09*****	−0.04	−0.01	**−0.08***	**−0.08****
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Lottery	**−0.20*****	**−0.26*****	**−0.11*****	0.02	**−0.28*****	**−0.29*****
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Young					−0.02	
					(0.05)	
Young×Queue					−0.01	
					(0.05)	
Young×Lottery					0.03	
					(0.05)	
High. Inc						0.03
						(0.03)
High. Inc×Queue						−0.02
						(0.05)
High. Inc×Lottery						0.08
						(0.05)
Num. obs.	1936	1924	1917	1908	1924	1924

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

The polling company employed a sampling strategy for this study that included weights. In our PAP, we had not envisioned the existence of weights because we had used a previous study by the same polling company which had been conducted on an internet sample. THe internet sample

Table O3: Experimental Results: Pooled Conditions

	Justness	Justness
Need-only + Queue (baseline)	0.761***	0.679***
	(0.012)	(0.058)
Lottery	−0.215***	−0.218***
	(0.023)	(0.023)
Num.Obs.	1924	1924

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

draws respondents form a pool of potential respondents for which the company already posses sociodemographic information, and therefore the sample is weighted ex-ante. In our current face-to-face mode of sampling, the sample is adjusted ex-post through weights.

Given the experimental nature of our study, we decided to keep the unweighted analysis in the main body of the paper as our main analysis, as it preserves the randomized nature of our study. In this section we replicate the analysed in Table 2 but include sample weights computed by the polling company. These weighted results are reported in Figure O4.

Table O4: Results with Sample Weights

	Main Effects				Het. Effects	
	Prog. Eval	Fairness	Efficiency	Favoritism	Just|Age	Just|Income
Need-only (baseline)	**0.90*****	**0.80*****	**0.69*****	**0.19*****	**0.82*****	**0.79*****
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Queue	**−0.04***	**−0.09*****	−0.04	−0.04	**−0.08***	**−0.08***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Lottery	**−0.18*****	**−0.24*****	**−0.10*****	−0.00	**−0.27*****	**−0.27*****
	(0.02)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)
Young					−0.05	
					(0.03)	
Young×Queueing					−0.01	
					(0.05)	
Young×Lottery					0.05	
					(0.05)	
High. Inc						0.04
						(0.03)
High. Inc×Queueing						−0.03
						(0.05)
High. Inc×Lottery						0.07
						(0.05)
Num. obs.	1936	1924	1917	1908	1924	1924

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Table O5: Experimental Results without MCMV eligible (all outcomes)

	Main Effects			
	Prog.Eval	Justness	Efficiency	Favoritism
Need (baseline)	**0.91*****	**0.82*****	**0.71*****	**0.18*****
	(0.01)	(0.02)	(0.02)	(0.02)
Queue	−0.03	**−0.10****	−0.05	−0.01
	(0.02)	(0.03)	(0.03)	(0.03)
Lottery	**−0.18*****	**−0.24*****	**−0.12*****	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
R^2	0.05	0.05	0.01	0.00
Adj. R^2	0.05	0.05	0.01	−0.00
Num. obs.	1174	1168	1163	1152
RMSE	0.36	0.45	0.47	0.39

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Appendix P Considerations about the Pre-analysis Plans

Appendix P.1 Quasi-experimental study pre-analysis plan

We registered a pre-analysis plan (PAP) related to hypotheses and research questions related to other manuscripts accepted or in preparation. We present, in this section, a discussion of how our analysis of the surveys discussed in Section 4 relates to this PAP and we have attached to the submission a redacted (and hence anonymous) copy of the PAP.

The main question related to lotteries' fairness we analyze in this manuscript was not part of our pre-analysis plan for the recent lotteries survey. But we included all the outcomes we study in this paper as part of our amended pre-analysis plan related to the early lotteries survey, even though they were not a central part of our pre-analysis plan. We focus our discussion regarding our PAP for the sections as sit directly relates to this manuscript.

Estimation Strategy We followed our pre-registered estimation strategy for our main analyses.

Attrition We followed our pre-registered protocol for examining and evaluating attrition.

Control Variables We pre-registered age, sex, race, formal employment and wages as our control variables for the 2011 lotteries study. In any case, our adjusted and unadjusted results are quite similar.

Hypotheses As mentioned, we did not include hypotheses related to lotteries' fairness and program evaluation in our pre-analysis for the recent lotteries' survey. But we had we expectations about early lotteries survey were:

"Hypothesis 12. The MCMV program is better evaluated by non-beneficiaries than by beneficiaries."

And then we wrote: "Somewhat counter-intuitively, we expect beneficiaries to have faced problems and difficulties with their new home, and therefore be more critical of the program than nonbeneficiaries, who have more idealized and less concrete perceptions."

We then expected to analyze this hypothesis using the questions: 1) Improved Lives, 2) Homeownership is a dream, 3) MCMV brings hope for a better life, 4) believe lotteries are fair, 5) would move. As stated in the hypotheses, we expected lotteries non-winners to have higher evaluations of the program than lottery winners. Yet results are largely null for these outcomes, except for perceptions of lotteries' fairness. Importantly, we expected a different result for the lotteries outcome in our early lotteries than the one we find in our paper.

The pre-analysis plan can be found here: https://osf.io/apd46/?view_only=

Appendix P.2 Experimental study pre-analysis plan

We pre-registered hypotheses, analysis, and code for our national survey experiment with approximately 2,000 respondents. The anonymous version of the PAP can be found here: https://osf.io/52r4x/?view_only=19fc7fddd25946acace3378dfe265762.

We present all the main and heterogeneous effects pre-registered in the main paper and we leave balance tests to the appendix. Our main deviation from the PAP is that we omit the pre-treatment control variable of class because that was not collected by the polling company (we had assumed it would be collected based on previous studies with this company, but this survey was conducted face-to-face with a shorter questionnaire). We also conducted unplanned additional analyses using the population weights provided by the survey company; we leave those weighted analysis to the appendix.

Results are overall consistent with our pre-registered expectations, except for support for the policy. We find mostly null results for the pre-registered heterogeneous effects, but our expectations about those were weaker than for the other hypotheses and we had pre-registered them as research questions.

Appendix Q Ethics Discussion

The research conducted in this paper was reviewed and deemed exempt by Yale Human Subject Committee (#2000020455) and approved by Emory's Human Subjects Committee (#IRB00101802) and FGV Ethics Committee (#04/2017).

Our survey with early lottery winners was partly conducted during the Covid-19 pandemic. All of our survey interviews in 2019 and 2020 were conducted via phone and enumerators were working from home. Our qualitative interviews were conducted in person in 2018. We collected subjects' names and contact information from public sources and via private vendors, respectively. This data collection was in accordance with national laws regarding privacy of personal information.

For our survey interviews, we obtained voluntary and informed consent via phone. Our enumerators read our consent form to subjects who then agreed or declined to answer our questionnaire. In addition to reading our consent form, we also offered subjects the option to receive the consent form via WhatsApp (a messaging application widely used in Brazil). That way, subjects would have our consent form in writing too. The consent form explicitly stated this questionnaire was part of a research project conducted by universities and subjects were also provided with a local research institution review board's contact information to facilitate access (via phone, email, or in person). We also trained enumerators to answer questions related to the nature of the research project to make it as clear as possible to subjects that we were not involved with any political group or governmental agency. Our survey did not engage in deception and we do not anticipate having intervened in the political process because the intervention analyzed here was conducted by the government (without collaboration from the authors).

For our qualitative MCMV interviews, we obtained oral consent in person and individuals were compensated for their time and reimburse for transportation costs. Interviews with TETO beneficiaries were reviewed by Human Subjects Research Review Boards. One of the authors of this

study was a principal investigator in a research partnership with TETO and this author had full access to the TETO interviews. Redacted and anonymized sections of the interviews were made available to the other authors.

Our experiment was embedded in a survey carried out in a nationally representative sample by Quaest, an accredited polling firm that is member of Brazil's trade association of polling companies and that abides by the association's practices and Brazil's fairly restrictive privacy laws. The study was also reviewed by FGV Ethics Committee (#161/2022) and by Emory's Human Subjects Committee (#IRB00004488) and, as specified in the approved protocols, an abbreviated oral consent form was read to all participants prior to all face-to-face interviews.

Appendix R Data Sources & Transparency

1. Administrative information on lotteries rolls in Rio de Janeiro were collected from public websites, converted to machine-readable formats, and compiled by the authors. City Hall shared with the authors the telegrams used to contact lottery winners and those were coded by a research assistant.

2. Survey data for the original survey (quasi-experimental analysis) was collected via telephone surveys conducted with company E-Field.

3. Qualitative interviews with MCMV beneficiaries were conducted via face-to-face interviews with support from company E-Field.

4. Face-to-face household interviews with a sample representative of the Brazilian voting-age population were with by Quaest

5. Additional data on *Minha Casa, Minha Vida* enrollments were obtained through requests using the Law of Access to Information.

6. Data on formal jobs and income from formal labor obtained from RAIS.

7. Cadunico (national registry data) data obtained via request to the Ministry of Social Development, which was subsequently merged into what is now the Ministry of Women, Family and Human Rights.