

On-line Appendix of the Article *Ride Your Luck! A Field Experiment on Lottery-based Incentives for Compliance*

Marco Fabbri*, Paolo Nicola Barbieri[†], Maria Bigoni[‡]

JEL-Classification: D04; H42

Keywords: Enforcement; Fare-dodging; Public Good; Randomized Control-Trial; Risk Attitudes; Sanctions.

**Principal Investigator and Corresponding author.* University of Amsterdam, School of Law, Nieuwe Achtergracht 166, Room A5-15, 1018WV Amsterdam, The Netherlands, E-mail: m.fabbri@uva.nl, Telephone: (+31) 024082843

[†]Centre for Health Economics, School of Business, Economics and Law, University of Gothenburg

[‡]Department of Economics, University of Bologna

Appendix A: Supplementary Material and Data Analysis

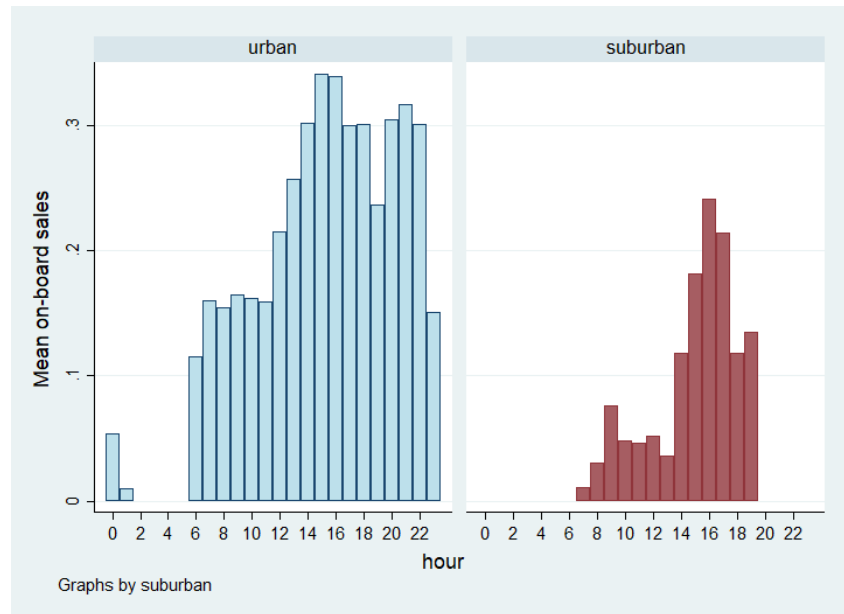


Figure A1: Average on-board ticket sales by time of day. Data include the whole sample of one-hour time slots during which buses circulated during the lottery period and during the preceding and following months. The left panel plots data relating to urban bus routes; the right panel reports data relating to suburban routes.

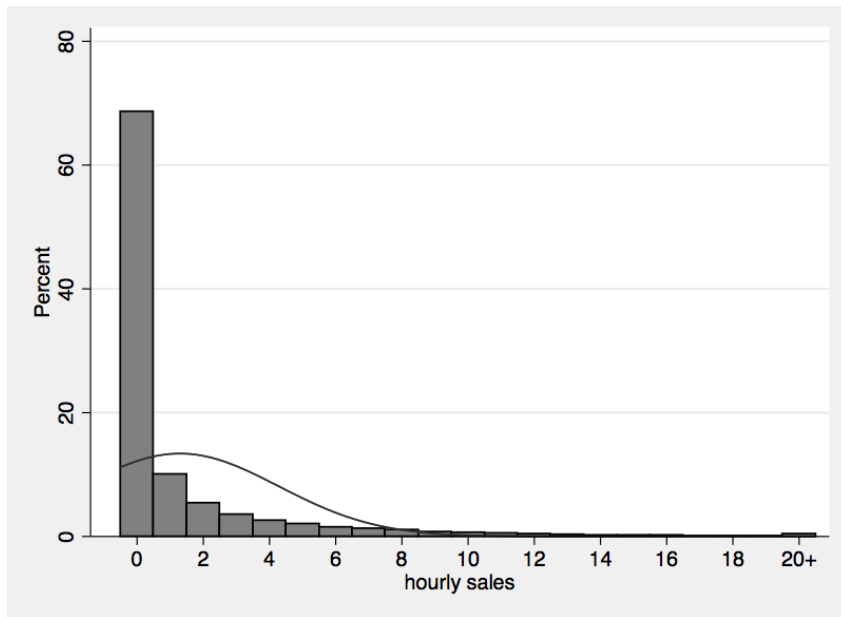


Figure A2: Distribution of on-board plus off-board sales per one-hour time slots. Data in our sample have a point mass at zero, indicating that in several slots buses were circulating although no tickets were sold. Notice that registering zero sales during a one-hour time slot does not imply that no passengers were using the bus during that period. For instance, passengers with weekly, monthly, or yearly tickets are not recorded in our data. Similarly, children under the age of 6, war veterans, company employees and their family members, people affected by disabilities, and those with an income below the poverty threshold automatically qualify to receive a pass for unlimited bus riding that does not need to be validated at each ride.

Table A1: Trends of *onboard sales* and *offboard sales* in the pre-lottery period

	Model 1	Model 2	Model 3	Model 4
Dep. variable: Sales	onboard	onboard	offboard	offboard
treated	35.518 (45.633)	27.104 (56.645)	-605.592 (493.498)	-628.993 (584.081)
day	-0.005** (0.002)	-0.006** (0.003)	-0.011 (0.018)	-0.014 (0.022)
treated × day	-0.002 (0.002)	-0.001 (0.003)	0.030 (0.025)	0.031 (0.029)
suburban=1		-116.032* (60.394)		-234.986 (466.848)
treated × suburban=1		31.117 (60.477)		89.099 (747.941)
suburban=1 × day		0.006* (0.003)		0.012 (0.023)
treated × suburban=1 × day		-0.002 (0.003)		-0.004 (0.037)
Constant	100.777** (48.633)	122.816** (59.249)	227.950 (353.755)	273.568 (435.219)
Time of the day f.e.	Yes	Yes	Yes	Yes
Line f.e.	Yes	Yes	Yes	Yes
N.obs.	22861	22861	22861	22861

Notes: Dependent variable for model 1 and 2: on-board tickets sales; dependent variable for models 3 and 4: off-board tickets sales. OLS regression, with line and hour of the day fixed effects. Standard errors robust for clustering at the line level. Model 2 and 4 isolate the effects of buses traveling respectively on urban routes and suburban routes. Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

Table A2: Total Number of Sanctions

	IRR	s.e.
treated	1.058	(0.050)
lottery	1.004	(0.044)
post-lottery	0.942	(0.042)
treated × lottery	1.060	(0.072)
treated × post-lottery	0.998	(0.072)
Constant	2.536***	(0.074)
N.obs.	2783	

Notes: Negative binomial regression. Dependent variable: total number of sanctions on a bus, in a day, during the lottery period and in preceding and following months. The coefficients report the incidence rate ratio. Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

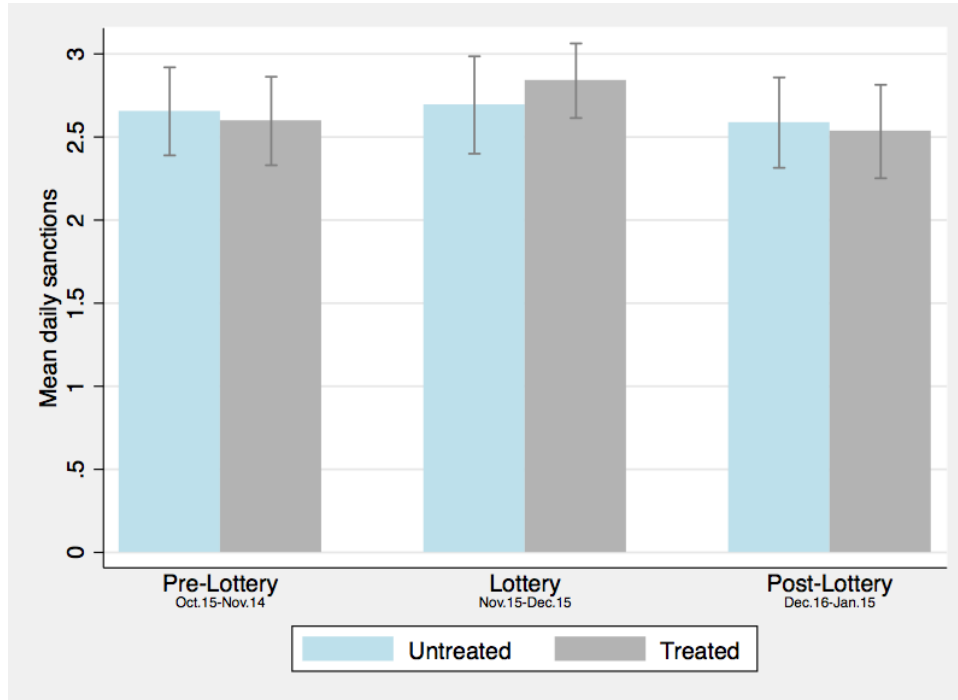


Figure A3: Average daily sanctions during the lottery period and in previous and following months in treated and control buses

Table A3: Lottery Effects on On-board Sales in the Initial and Final Periods of Implementation

	Model 1	Model 2	Model 3	Model 4
treated	0.077*** (0.027)	0.071** (0.026)	0.032 (0.079)	0.027 (0.080)
lottery	-0.111 (0.078)	-0.186** (0.085)	-0.694*** (0.151)	-0.547*** (0.148)
post-lottery	-0.275*** (0.088)	-0.279*** (0.088)	-0.627*** (0.141)	-0.644*** (0.144)
treated × lottery	0.058*** (0.016)	0.076*** (0.015)	0.140*** (0.033)	0.110** (0.052)
treated × post-lottery	0.053** (0.020)	0.054** (0.021)	0.083 (0.054)	0.087 (0.059)
Constant	-0.142 (0.111)	-0.163 (0.117)	0.145 (0.184)	0.120 (0.288)
Time of the day f.e.	Yes	Yes	Yes	Yes
Line f.e.	Yes	Yes	Yes	Yes
Day f.e.	Yes	Yes	Yes	Yes
N.obs.	55226	53986	16355	17413

Notes: Dependent variable: on-board ticket sales. Linear regressions, with line, time of day, and day fixed effects. Standard errors robust for clustering at the line level. Models 1 and 3 exclude from the sample the second half of the lottery period, Models 2 and 4 exclude from the sample the first half of the lottery period, Models 3 and 4 consider the restricted sample of hourly time slots during which at least one ticket was sold, either on or off-board. Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

dal 15 Novembre
al 15 Dicembre 2014

GRANDE CONCORSO Compra, Viaggia e Vinci...

Acquistando il biglietto del bus dalle emettitrici automatiche di bordo (EMY)[®], parteciperai all'estrazione di biglietti vincenti che daranno diritto a

BUONI SPESA DI € 500

da utilizzare presso i supermercati di Coop Adriatica, nelle sedi di Rimini.
Cosa aspetti? Sali a bordo, acquista e vinci!

* Il concorso è riservato ai possessori di biglietti bus acquistati durante il periodo di validità del concorso dalle emettitrici automatiche EMY, poste a bordo di alcuni bus Start Romagna del bacino di Rimini contrassegnati da apposita segnaletica interna. Regolamento completo su www.startromagna.it e www.amrimini.it

servizi@startromagna.it www.startromagna.it

COME SI PARTECIPA

La partecipazione al concorso è riservata ai possessori di biglietti acquistati nel periodo 15 novembre 2014 - 15 dicembre 2014 dalle emettitrici EMY presenti a bordo dei bus START ROMAGNA del bacino di Rimini individuati da apposita segnaletica interna (per un totale di 50 autobus). I partecipanti devono conservare il biglietto originale acquistato nel periodo di validità del concorso.

I PREMI

Sono previsti n. 6 premi in buoni spesa del valore di euro 500 ciascuno da utilizzarsi presso i supermercati della Coop Adriatica - sedi di Rimini.

ESTRAZIONE PREMI

Ogni biglietto acquistato da EMY è contraddistinto da un numero identificativo unico, riportato sul fronte del biglietto.

Sono previste tre diverse estrazioni, riferite ai biglietti acquistati nei periodi:

1. dal 15 al 21 novembre 2014
2. dal 22 novembre al 1 dicembre 2014
3. dal 2 al 15 dicembre 2014



Per ogni periodo sopra indicato saranno estratti i numeri identificativi di due biglietti vincenti; le estrazioni saranno effettuate presso la sede di Agenzia Mobilità alla presenza del funzionario della Camera di Commercio di Rimini entro la settimana successiva a quella di riferimento.

ESEMPIO
DI COMPOSIZIONE
VINCENTE

PUBBLICIZZAZIONE NUMERI VINCENTI

I numeri dei biglietti estratti e le modalità di ritiro dei premi saranno resi noti a mezzo stampa, con avvisi sui bus e sui siti internet di START ROMAGNA (www.startromagna.it) e Agenzia Mobilità (www.amrimini.it).

Figure A4: An example of a flyer advertising the lottery. The original text in Italian states: (Left panel, upper part) *From November 15 to December 15, 2014, the 'Buy, Travel, and Win...' contest;* (Left panel, lower part) *By purchasing an on-board bus ticket from the EMY ticket-selling machines, you will take part in a lottery contest awarding prizes worth €500 in vouchers to be used at Coop Adriatica stores. What are you waiting for? Jump on-board, buy the ticket, and win!. The contest is open to holders of on-board tickets purchased from the automatic EMY ticket-selling machines on some of the Start Romagna buses serving the municipal area of Rimini that are marked by specific signage inside the vehicles. See the complete guidelines at www.startromagna.it or www.amrimini.it;* (Right panel) *How to participate: participation is reserved for holders of on-board tickets purchased from EMY ticket-selling automatic machines between November 15 and December 15, 2014 on some of the Start Romagna buses (a total of 54 buses) serving the municipal area of Rimini that are marked by specific signage placed inside the vehicles. Participants must keep the ticket purchased during the contest. Prizes: 6 vouchers worth €500 each for shopping at Coop Adriatica supermarkets. Lottery: Each ticket is identified by a unique serial number, printed on the front side of the ticket. Three lottery rounds will take place: November 15-21; 2) November 22-December 1; 3) December 2-15). For each period, two winning tickets will be drawn. Lottery draws will take place within one week after the end of the respective period at Agenzia Mobilità's headquarters, and a representative of the Chamber of Commerce will be present. Publication of Winning Ticket Numbers: Winning ticket numbers will be publicized via the media and a press release, publication on buses, and posts on Start Romagna and Agenzia Mobilità websites.*

Appendix B: Robustness Check

In Table A4 we replicate the analysis reported in Table ?? by using a Poisson regression in order to take into account the count nature of the on-board sales data. As in Table ??, models 1 and 2 consider respectively the whole sample of one-hour time slots and the restricted sample including only slots when at least one ticket was sold, and models 3 and 4 replicate respectively models 1 and 2 additionally including the dummy for suburban bus routes. To facilitate the interpretation of the regression coefficients, we report the incidence rate ratio (that is, we exponentiate the Poisson regression coefficient): the estimated effect of the incentive scheme results from the expected on-board sales count multiplied by the reported coefficient of the interaction term *treated * lottery*.¹ Therefore, a coefficient of *treated * lottery* between 0 and 1 implies a decrease in on-board sales, and a coefficient larger than 1 implies an increase in the count of the outcome variable.

The coefficient of the interaction term *treated * lottery* is statistically significant at the conventional level or better in any model specification, suggesting that the incentive scheme determined a 16%-17% increase in on-board sales. Models 3 and 4 in Table A4 suggest that the increase in on-board sales is driven by buses operating in urban routes, while passengers riding buses operating on suburban routes seem to have responded significantly less (at least in model 3, where the whole sample of hourly slots is considered) to the lottery introduction.

As an additional robustness check, Table A5 in Appendix A reports the results of negative binomial regression and zero-inflated negative binomial regression models in order to account for overdispersion and to consider the point mass at zero. The zero-inflated negative binomial regression models the data generating process as a combination of two processes – the first involving the decision whether to use the bus, the second involving the decision whether to purchase the ticket. In models 1 and 3 of Table A5 we assumed that people’s decision to catch a bus is influenced by weather conditions (proxied by mean daily temperature). Again, the results of these alternative model specifications are qualitatively and quantitatively similar to those presented above.

Similarly, as a robustness check we also replicate the analysis on off-board tickets reported

¹The incidence rate ratio is a common way to interpret the estimation results of Poisson or Negative Binomial models. Regression coefficients of these models report the difference between the log of expected counts. Formally, this can be written as $\beta = \lg(\mu x_0 + 1) - \lg(\mu x_0)$, where β is the regression coefficient, μ is the expected count and the subscripts represent where the predictor variable, say x , is evaluated at x_0 and $x_0 + 1$ (implying a one unit change in the predictor variable x). Given that the difference of two logs is equal to the log of their quotient, $\lg(\mu x_0 + 1) - \lg(\mu x_0) = \lg\left(\frac{\mu x_0 + 1}{\mu x_0}\right)$, and therefore, we could have also interpreted the parameter estimate as the log of the ratio of expected counts: this explains the “ratio” in incidence rate ratios. In addition, what we referred to as a count is technically a rate. A rate is defined as the number of events per time. Our response variable is the number of tickets sold in a hourly slot which, by definition, is a rate. Hence, we could also interpret the regression coefficients as the log of the rate ratio: this explains the “rate” in incidence rate ratio. Finally, the rate at which events occur is called the incidence rate; thus we arrive at being able to interpret the coefficients in terms of incidence rate ratios.

in Table ?? by implementing a Poisson regression. As before, we include day, hour of the day, and bus-route fixed effects, we cluster standard errors by bus-route, we consider restricted samples (model 2 and 4), and we disentangle the effects on urban and suburban routes (model 3 and 4). The coefficients in Table A6 report the incidence rate ratio. The results confirm previous findings regarding the effects of the lottery on off-board sales. The coefficient of the interaction term $treated \times lottery$ in models 1 and 2 is statistically not significant, indicating that overall the incentive scheme did not affect off-board sales. In model 3 and 4 the coefficient of $treated \times lottery$, which isolates the effect on off-board sales for urban routes, is once again not statistically different from zero. Conversely, the coefficient $treated \times lottery \times suburban$ is smaller than one and significant at the conventional level or better, suggesting that on suburban routes the introduction of the lottery caused a 39% - 24% reduction of off-board ticket sales.

Table A4: On-board ticket sales, Poisson regression

	Model 1	Model 2	Model 3	Model 4
treated	1.483*** (0.225)	1.095 (0.132)	1.415** (0.206)	1.094 (0.133)
lottery	0.234*** (0.059)	0.334*** (0.074)	0.224*** (0.054)	0.325*** (0.072)
post-lottery	0.107*** (0.033)	0.287*** (0.076)	0.102*** (0.031)	0.279*** (0.074)
treated × lottery	1.158** (0.075)	1.174*** (0.062)	1.214*** (0.071)	1.176*** (0.062)
treated × post-lottery	1.067 (0.104)	1.105 (0.089)	1.137** (0.074)	1.113 (0.092)
suburban=1			0.000*** (0.000)	0.000*** (0.000)
treated × suburban=1			2.322*** (0.466)	1.241 (0.500)
lottery × suburban=1			2.013*** (0.394)	1.422** (0.230)
post-lottery × suburban=1			2.077*** (0.223)	1.535 (0.410)
treated × lottery × suburban=1			0.423*** (0.106)	0.796 (0.212)
treated × post-lottery × suburban=1			0.331*** (0.053)	0.699 (0.278)
Constant	0.009*** (0.009)	0.079** (0.082)	0.010*** (0.009)	0.090** (0.099)
Time of the day f.e.	Yes	Yes	Yes	Yes
Line f.e.	Yes	Yes	Yes	Yes
Day f.e.	Yes	Yes	Yes	Yes
N.obs.	64714	20284	64714	20284

Notes: Dependent variable: on-board ticket sales. Poisson regressions, with line, time of day, and day fixed effects. The coefficients report the incidence rate ratio. Standard errors robust for clustering at the line level. Model 1 includes the full sample, Model 2 considers the restricted sample of hourly time slots during which at least one ticket was sold, either on or off-board, Model 3 isolates the effects of buses traveling respectively on urban routes and suburban routes, Model 4 isolates the effects of buses traveling respectively on urban routes and suburban routes and considers the restricted sample of hourly time slots during which at least one ticket was sold. Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

Table A5: On-board sales, zero-inflated negative binomial regressions and negative binomial regressions

	Model 1	Model 2	Model 3	Model 4
treated	1.623*** (0.248)	1.100 (0.141)	1.539*** (0.230)	1.099 (0.142)
lottery	0.223*** (0.085)	0.322*** (0.084)	0.211*** (0.082)	0.314*** (0.082)
post-lottery	0.098*** (0.043)	0.271*** (0.083)	0.092*** (0.041)	0.262*** (0.081)
treated × lottery	1.124* (0.079)	1.176*** (0.066)	1.190*** (0.079)	1.178*** (0.067)
treated × post-lottery	1.050 (0.114)	1.126 (0.095)	1.134 (0.088)	1.137 (0.098)
suburban=1			0.000*** (0.000)	0.000*** (0.000)
treated × suburban=1			2.174*** (0.487)	1.253 (0.508)
lottery × suburban=1			1.811*** (0.384)	1.407** (0.235)
post-lottery × suburban=1			2.023*** (0.268)	1.606* (0.399)
treated × lottery × suburban=1			0.431*** (0.131)	0.799 (0.221)
treated × post-lottery × suburban=1			0.323*** (0.069)	0.668 (0.259)
Constant	0.011*** (0.011)	0.082** (0.086)	0.011*** (0.011)	0.094** (0.104)
Time of the day f.e.	Yes	Yes	Yes	Yes
Line f.e.	Yes	Yes	Yes	Yes
Day f.e.	Yes	Yes	Yes	Yes
N.obs.	64714	20284	64714	20284

Notes: Dependent variable: on-board ticket sales. Zero-inflated negative binomial (Models 1 and 3) and negative binomial (Models 2 and 4) regressions, with line, time of day, and day fixed effects. The zero-inflated negative binomial models assume that the first process is influenced by weather conditions (daily mean temperature). The coefficients report the incidence rate ratio. Standard errors robust for clustering at the line level. Model 1 includes the full sample, Models 2 considers the restricted sample of hourly time slots during which at least one ticket was sold, either on or off-board, Model 3 isolates the effects of buses traveling respectively on urban routes and suburban routes, Models 4 isolates the effects of buses traveling respectively on urban routes and suburban routes and considers the restricted sample of hourly time slots during which at least one ticket was sold. Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

Table A6: Off-board ticket sales, Poisson regression

	Model 1	Model 2	Model 3	Model 4
treated	1.288** (0.149)	0.988 (0.079)	1.281** (0.160)	0.987 (0.080)
lottery	0.641** (0.140)	0.755 (0.165)	0.639** (0.135)	0.745 (0.162)
post-lottery	0.425*** (0.102)	0.802 (0.180)	0.431*** (0.102)	0.804 (0.180)
treated × lottery	0.931 (0.062)	0.940 (0.058)	0.964 (0.060)	0.958 (0.056)
treated × post-lottery	0.851 (0.141)	0.876 (0.090)	0.864 (0.155)	0.875 (0.093)
suburban=1			0.367*** (0.078)	0.763** (0.086)
treated × suburban=1			0.994 (0.344)	0.971 (0.127)
lottery × suburban=1			1.104 (0.142)	1.156* (0.086)
post-lottery × suburban=1			0.838 (0.272)	0.969 (0.117)
treated × lottery × suburban=1			0.609*** (0.108)	0.760** (0.090)
treated × post-lottery × suburban=1			0.802 (0.221)	1.021 (0.171)
Constant	0.509 (0.296)	5.168*** (2.568)	0.466 (0.281)	4.919*** (2.520)
Time of the day f.e.	Yes	Yes	Yes	Yes
Line f.e.	Yes	Yes	Yes	Yes
Day f.e.	Yes	Yes	Yes	Yes
N.obs.	64714	20284	64714	20284

Notes: Dependent variable: off-board ticket sales. Poisson regressions, with line, time of day, and day fixed effects. The coefficients report the Incidence Rate Ratio. Standard errors robust for clustering at the line level. Model 1 includes the full sample, Model 2 considers the restricted sample of hourly time slots during which at least one ticket was sold, either on or off-board, Model 3 isolates the effects of buses traveling respectively on urban routes and suburban routes, Model 4 isolates the effects of buses traveling respectively on urban routes and suburban routes and considers the restricted sample of hourly time slots during which at least one ticket was sold. Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

Appendix C: Theoretical Model

Suppose that the economy consists of a continuum of N bus riders with constant absolute risk aversion (CARA) utility given by $u(Y, i) = -e^{-iY}/i$, $i \neq 0$, where Y denotes consumption and i is a parameter denoting risk preferences.² We assume that bus riders differ only in their degree of risk aversion, i , which is distributed in the interval (\underline{i}, \bar{i}) according to a cumulative density function $G(i)$, where riders with $i > 0$ are risk-averse and riders with $i < 0$ are risk-loving.

Every bus rider i faces the binary decision of whether to pay the travel fare t or to free ride. A consumer who does not pay the travel fare faces the risk of being sanctioned, in which case he will have to pay an exogenous fine $F \in (t, Y)$. We assume that the probability of being sanctioned π is exogenous and independent from the number of paying consumers and/or free riders.

A bus rider i pays the travel fare t if

$$u(Y - t, i) \geq \pi u(Y - F, i) + (1 - \pi)u(Y, i) \quad (1)$$

or

$$1 \geq \frac{u(Y, i) - u(Y - t, i)}{\pi[u(Y, i) - u(Y - F, i)]} \quad (2)$$

Under the stated assumptions regarding the utility function, we can write inequality (2) as

$$1 \geq \frac{-e^{-iY} + e^{-i(Y-t)}}{\pi[-e^{-iY} + e^{-i(Y-F)}]} = \frac{e^{it} - 1}{\pi(e^{iF} - 1)} \quad (3)$$

Proposition 1. *For any given probability π , fare t , and fine F , there will be a marginal rider \hat{i} who is indifferent between paying the fare and fare-dodging, and all the riders with $i > \hat{i}$ will pay the fare, while all those with $i < \hat{i}$ will free ride.*

Proof. Differentiating the right hand side (RHS) of (3), for i we get

$$\frac{(t - F)e^{i(F+t)} + Fe^{Fi} - te^{it}}{\pi(e^{Fi} - 1)^2} \quad (4)$$

which is negative if

$$Fe^{iF}(1 - e^{it}) - te^{it}(1 - e^{iF}) < 0.$$

²The assumption is consistent with the literature; see, for example, Lange et al. (2007). We further assume that $U(Y, i) = Y$ if $i = 0$.

This implies that

$$\begin{cases} \frac{Fe^{iF}}{1-e^{iF}} < \frac{te^{it}}{1-e^{it}} & \text{if } i < 0 \\ \frac{Fe^{iF}}{e^{iF}-1} > \frac{te^{it}}{e^{it}-1} & \text{if } i > 0 \end{cases}$$

which is always true $\forall F > t$. In addition, $\lim_{i \rightarrow -\infty} \frac{e^{it}-1}{\pi(e^{iF}-1)} = \frac{1}{\pi} > 1$ and $\lim_{i \rightarrow \infty} \frac{e^{it}-1}{\pi(e^{iF}-1)} = 0$. \blacksquare

Adding a carrot to the stick. Suppose that in order to reduce fare-dodging the local transportation authority introduces a lottery. Agents purchasing a ticket participate in the lottery and receive with probability q a reward equal to R . Under the lottery rules, the probability of agent i winning prize R is inversely proportional to the total number N_p of agents who pay the fare: $q = \frac{1}{N_p}$.

Proposition 2. *If, absent the lottery, the proportion of free riders is strictly higher than 0, and strictly lower than 1, for any given probability π , fare t , fine F , and reward $R > 0$, the introduction of the lottery cannot increase, and possibly decreases, the proportion of free riders.*

Proof. A bus rider i pays the travel fare t if

$$qu(Y - t + R, i) + (1 - q)u(Y - t, i) \geq \pi(u(Y - F, i)) + (1 - \pi)u(Y, i) \quad (5)$$

or

$$1 \geq \frac{u(Y, i) - u(Y - t, i)}{\pi[u(Y, i) - u(Y - F, i)]} - \frac{q(u(Y - t + R, i) - u(Y - t, i))}{\pi[u(Y, i) - u(Y - F, i)]} \quad (6)$$

With a CARA utility function, Equation (5) yields:

$$1 \geq \frac{e^{it} - 1}{\pi(e^{iF} - 1)} - q \times \frac{(e^{iR} - 1)e^{i(t-R)}}{\pi(e^{iF} - 1)} \quad (7)$$

Since the second element of the RHS of Inequality (7) is always positive, the introduction of the lottery cannot increase the level of i that equals the two sides of (7), thus possibly inducing more agents to pay the fare.³ However, the positive impact of the lottery on compliance is attenuated by the fact that the probability of winning the prize decreases in the number of compliers, which implies that the second element of the RHS of inequality (7) can be non-monotonic in i , depending also on $G(i)$.

A numerical example will show that it is possible to introduce a “self-financed” lottery that increases fare revenues so as to completely cover the cost of the prize.

Numerical example. Consider a population of riders of size N , with two types of agents, who differ in their risk attitudes. A proportion $f = G(\underline{i})$ of the agents are moderately risk-loving with preferences characterized by a parameter $\underline{i} < 0$, while the others are risk-averse ($\bar{i} > 0$).

³Whether the fraction of compliers acually increases depends on the distribution of preferences $G(i)$.

Let the probability of detection be $\pi = \frac{1}{N}$, and let the sanction be equal to $F = t \times f \times N$. Under these conditions, it is straightforward to see that a risk-neutral agent would not pay the fare, as

$$Y - t < \pi(Y - F) + (1 - \pi)Y$$

Let us assume, instead, that \bar{i} is such that condition (3) is satisfied, that is,

$$1 \geq \frac{e^{\bar{i}t} - 1}{\pi(e^{\bar{i}F} - 1)} = \frac{N(e^{\bar{i}t} - 1)}{e^{\bar{i}t}fN - 1}$$

This implies that, absent any lottery, a fraction f of the agents free rides, while the remaining $N(1 - f)$ pay the travel fare.

Let us now consider the introduction of a lottery with a prize $R = F$, where the probability of getting the prize for an agent who pays the fare is equal to 1 over the total number of compliant riders in the population. The introduction of the lottery simply strengthens the incentive to pay the fare; therefore all the risk-averse agents will continue to comply under the new regime. The risk-loving agents instead will comply if condition (7) is satisfied when all agents pay the fare, that is, for $q = \frac{1}{N}$. This implies that:

$$1 \geq \frac{e^{it} - 1}{\pi(e^{iF} - 1)} - e^{i(t-F)} = N \frac{e^{it} - 1}{(e^{it}fN - 1)} - e^{it(1-fN)} = N \frac{e^{it} - 1}{(e^{it}fN - 1)} - \frac{e^{it}}{e^{it}fN}$$

Because

$$\lim_{i \rightarrow -\infty} N \frac{e^{it} - 1}{(e^{it}fN - 1)} - \frac{e^{it}}{e^{it}fN} = -\infty$$

there must be a value i^* such that $1 \geq \frac{e^{it} - 1}{(e^{it}fN - 1)} - \frac{e^{it}}{e^{it}fN}$ for all $i < i^*$. Therefore, if $i < i^*$, all agents will pay the fare after the introduction of the lottery, and the total amount of money collected through the fares will increase by $f \times N \times t$, which by assumption is exactly enough to cover the cost of the prize R .

As a numerical example with parameters that are roughly consistent with those in our field experiment, suppose that the number of travelers is $N = 1000$, the travel fare is $t = 2$, and the sanction is $F = 80$. Let $\underline{i} \leq 0.1$, $\bar{i} \geq 0.1$, and $f = \frac{1}{25}$. Without any lottery, only the risk-averse agents pay the fare, while the risk-loving free ride. Figure B1 illustrates the LHS and RHS of inequality (3), with the selected parametrization.

When a lottery is introduced, all agents prefer to pay the fare, the risk-averse ones because they are afraid of the sanction and the risk-loving ones because they are attracted by the prize. Figure B2 illustrates the LHS and RHS of inequality (7), with the selected parametrization. The number of riders who pay the fine increases by $f \times N = 40$ as a consequence of the introduction of the lottery; therefore the revenues increase by an amount that is sufficient to cover the cost

of the prize, equal to 80.⁴

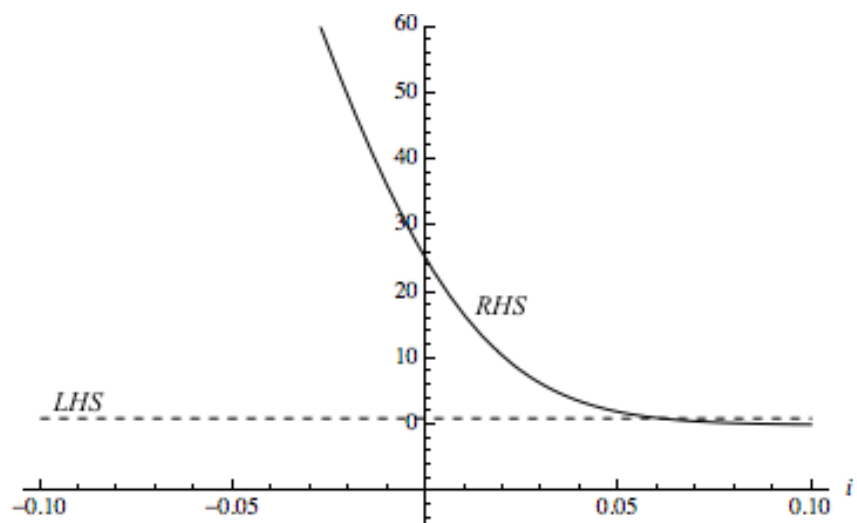


Figure B1: Condition for compliance without the lottery (inequality 3)

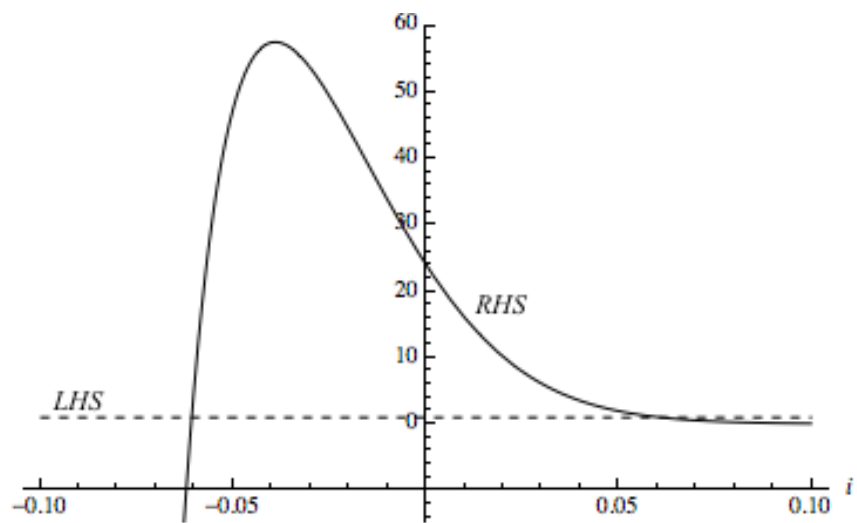


Figure B2: Condition for compliance with the lottery (inequality 7)

⁴In the field experiment, the prize is in fact roughly ten times larger than the fine, but the share of free riders is also about ten times larger than in this example. This implies that if the reward is large enough to convince all free riders to comply, the lottery will indeed generate an increase in the revenues which covers the cost of the prize.