High School Teens and Museums Attendance: A Field Experiment

Patrizia Lattarulo, IRPET – Tuscany's Regional Institute for Economic Planning, Villa La Quiete alle Montalve, Via Pietro Dazzi, 1 - 50141 Firenze, Italy, patrizia.lattarulo@irpet.it

Marco Mariani, IRPET – Tuscany's Regional Institute for Economic Planning, Villa La Quiete alle Montalve, Via Pietro Dazzi, 1 - 50141 Firenze, Italy, marco.mariani@irpet.it

Laura Razzolini, Department of Economics, School of Business, Virginia Commonwealth University,301 West Main Street, Box 4000, Richmond, VA23284, USA, lrazzolini@vcu.edu

August 18, 2015

Abstract

This paper reports results from a field experiment conducted to study incentives offered to high school teens to motivate them to visit art museums. A vast literature has been developed to study the design of incentives in firms and consumers, but not much is known about incentives to young adolescents to perform a task with positive educational spillovers. Students in the control group receive a flier containing basic information and opening hours of a main museum in Florence, Italy – Palazzo Vecchio. Students in the first encouragement group receive the flier and a short presentation about the exhibit conducted by an art expert; students in the second encouragement arm, in addition to the flier and the presentation receive also a non-financial reward in the form of extra-credit points towards their school grade. The analysis yields two main findings. First, the extra-credit, non-financial reward, is more effective at inducing the students to undertake the encouraged visit than either the simple presentation or the basic information with the flier. Second, in a longer time horizon the extra-credit reward does not induce a change in behavior with respect to the simple presentation.

JEL Classifications: M52, D82, C93, I28, Z10, Z18
Keywords: incentives, non-monetary rewards, cultural consumption, cluster-randomized experiment, randomization inference

* Corresponding author: Marco Mariani, marco.mariani@irpet.it

1. Introduction

Adolescents and teens often visit museums with their families or school teachers and classmates, but when it's up to them to choose how to allocate their free time, museums rarely appear at the top of their preferences (Gray, 1998).This phenomenon realizes in spite of the 'open doors' policy followed by most museums and their attempts to portray an image of educational and entertaining institutions. The literature on cultural consumption reports a correlation between participation to cultural events and household income and status (e.g. Falk and Katz-Gerro, 2015). Very little attention, however, has been paid to incentives in cultural organizations, such as museums and art galleries, to pursue the goals of increasing attendance and visits, particularly by adolescents and teens'. Our paper begins to fill this gap by providing evidence from a field experiment designed to study incentives offered to high school teens to motivate them to visit an art museum in Florence, Italy. The aim of the experiment is to encourage individual museum attendance during the high school years, and to identify best practices to transform this behavior into a long run cultural consumption.

The field experiment was conducted in Florence, Italy during the spring of 2014. The experiment was designed to identify the best incentives to offer to high school teens to motivate them to visit an art museum. Students in the control group received a flier containing basic information and opening hours of a main museum in Florence – Palazzo Vecchio. Students in the first encouragement group received the flier and a short presentation about the exhibit by an art expert from the museum; students in the second encouragement group, in addition to the flier and the presentation, received a non-financial reward in the form of extra-credit points towards their school grade.

Field experiments have been increasingly popular in economic analyses (see Levitt and List, 2009 for a general survey). However, their application to this area of research is novel. Recently published contributions in the field of cultural economics by Suárez-Vázquez (2011), Barkshi and Thorsby (2014), Berlin et al. (2015) report results from laboratory or field experiments, but none focuses on museum attendance and means to increase its practice. In our field experiment, about 300 high school students from 15 different classes were offered one of three different levels of encouragement inviting them to visit the museum for free during a given period of time. Students were also asked to fill in a questionnaire about their (and their families') background characteristics and attitudes toward museum attendance and cultural consumption in general. Assignment to the three levels of encouragement was randomized and occurred at the level of classes, which configures our experiment as a clustered randomized trial (Frangakis et al. 2002; Duflo et al. 2007). The students were asked to prove their visit to the museum by returning the admission ticket. Finally, six months later, the students were contacted

again and information about any additional number of museums visits done since the intervention was collected.

The aim of this research is to estimate which form of encouragement is associated with a higher probability of visiting the museum, and to establish whether and how each encouragement regime affects the students' subsequent behavior. Randomization inference techniques (see Gerber and Green, 2012 and Imbens and Rubin, 2015) were used to perform the analysis.

Results suggest that classroom visits and communication by museum personnel are very likely to increase future voluntary museum attendance, independent on whether the students did or did not undertake the encouraged visit at the time of the intervention. The additional provision of extra-credit points is shown to boost immediate execution of the encouraged visit, but has limited effects on later behavior.

The paper is organized as follows. Section 2 describes the experimental design and presents the variables of interest. Section 3 introduces the randomization inference approach for estimating and testing these variables, while Section 4 reports the results of the analysis. Section 5 concludes the paper.

2. Experimental Design

The field experiment was run in Florence, Italy, during the Spring of 2015. Students attending local high school schools were offered the opportunity to visit Palazzo Vecchio. Palazzo Vecchio houses the city's main offices and is one of the most visited museums in Florence. It is located in the city center and is easily reachable from different parts of the city. The entrance is free to individuals 18 and younger.[1]

The experiment consisted in using different types of encouragement to induce the high school students to visit the proposed museum: a flier containing basic information, a presentation by an art expert about the museum, and a non-financial reward in the form of extra-credit toward the final school grade. Different classes were randomly assigned to different types of encouragement.

The experiment involved 297 students from 15 different classes from three different high schools. All students attended the 4th year of high school and were aged 17-18. All high schools offered a same program of studies involving a mix of humanities, mathematics and scientific subjects. This type of high school provides no vocational education and its main purpose is to prepare students for college or university. All high schools are located in Florence and are attended by students from a similar socio-economic background.

In the experiment, each class was randomly assigned to one of three different levels of encouragement W:

[1]For students over 18 years old, a free ticket was provided, to guarantee an equal treatment.

- *W = f*: students receive a flier containing basic information, as opening hours, a brief description of the museum;
- *W = p:* in addition to the previous, students receive a short presentation about the exhibit conducted by an art expert;[2]
- *W = r*: in addition to the previous two, students receive a non-financial reward in the form of extra-credit points towards their final school grade.[3]

Table 1 – Experimental Design.

	N. of classes	*N. of Students (%)*
FLIER	5	93 (32%)
FLIER & PRESENTATION	5	97 (33%)
FLIER & PRESENTATION & REWARD	5	104 (35%)
Total	15	297 (100%)

Table 1 shows the number of classes and students assigned to each encouragement treatment. The randomized assignment was done by clusters at the class level. Students and teachers were told that they were participating in a survey on teen's cultural consumption and were not informed that they were part of an experiment with different forms of encouragement.[4]

Students were also asked to complete a survey at three different points in time: when the experiment was administered, to collect general information about individual characteristics - habits of cultural and leisure consumption for the students and their parents, school performance, friendship ties in and out of the class; after two month, to assess whether a visit to the museum happened, by collecting the admission tickets to the museum; after 6 months, to assess future voluntary attendance to museums after the experiment. Table 2 reports that timeline for the field experiment.

Table 2– Timeline.

When	*W = flier*	*W = presentation*	*W = reward*
	Students are informed that they will be involved in a study about cultural consumption. A **flier** of Palazzo Vecchio with opening hours is distributed.		

[2]The presentation, a form of motivational encouragement, consisted in a talk by an art expert from Palazzo Vecchio, specialized in communicating to a juvenile audience. The presentation followed a strict protocol, with moments of interaction scheduled and used the support of materials and illustrations.

[3]The reward encouragement consisted of extra-credit points toward the final class grade, specifically from 0.5 to 1.5 extra points on a 1-10 scale.

[4]The randomization occurred between pairs of classes located in different schools, or different buildings in a same school or floors in a same building.

<table>
<tr><td rowspan="4">First visit,
late March/
early April 2014</td><td></td><td colspan="2">A museum operator talks about Palazzo Vecchio for 15 minutes</td></tr>
<tr><td></td><td></td><td>Students are told that the visit to PalazzoVecchio will be rewarded with extra credit</td></tr>
<tr><td colspan="3">Students complete a questionnaire about their background characteristics, cultural consumption habits and details about within-classroom friendship ties</td></tr>
<tr><td colspan="3">Students who visit Palazzo Vecchio within two months are asked to bring the entry ticket back to their teacher</td></tr>
<tr><td>Second visit,
After two months</td><td colspan="3">Collect entry tickets</td></tr>
<tr><td>Third visit,
After six months</td><td colspan="3">Students complete a questionnaire about the number of individual visits to museums done in town and out of town in the last six months</td></tr>
</table>

Table 3 describes the main characteristics of the students in the sample by treatment received. Overall, 35% of the students are male, only a minority is born abroad (6.1%), 42% lives in the surroundings of the town, outside the city center. The majority of students are 17 years old, with only a few older (5.8%).

Table 3 –Descriptive statistics about students

Information about students	OVERALL	FLIER	PRESENT.	REWARD
Male (1/0)	0.354	0.215	0.299	0.529
Born abroad (1/0)	0.061	0.022	0.093	0.067
Suburban resident(1/0)	0.422	0.387	0.320	0.548
Older than cohort (1/0)	0.058	0.075	0.062	0.039
Information about the high School				
Most friends are classmates (1/0)	0.486	0.452	0.474	0.529
N. classmates who are friends	3.582	4.441	2.635	3.773
N. classmates who visited the museum	0.565	0.237	0.412	1.000
Current GPA (out of 10)	6.817	6.720	6.789	6.931
Interested only in humanities (1/0)	0.262	0.226	0.392	0.173
Leisure & Cultural habits				
Interested in politics (1/0)	0.228	0.215	0.206	0.260
Volunteer (1/0)	0.167	0.204	0.144	0.154
N. museum visits during last year	3.867	3.269	4.742	3.587
Visited Palazzo Vecchio previously (1/0)	0.721	0.677	0.763	0.721
Visited Cappella Brancacci previously (1/0)	0.214	0.258	0.330	0.067
Visited SM Novella museum previously (1/0)	0.337	0.290	0.392	0.327
Parents' education				
At least one parent employed	0.157	0.151	0.175	0.144
Both high school degree	0.327	0.376	0.247	0.356

At least one has college degree	0.241	0.258	0.258	0.212
How often parents go to museums	0.197	0.183	0.247	0.163

Average reported students' GPA is 6.8 out of 10 points and around 26% of the sample considers him/herself keen on humanities. Around half of the students report having the majority of friends in the same class (45%), and each student lists about having close 3.5 friends in the class. More than half of the friends listed by each students does the encouraged visit (56%). Each student (on average) visited more than three museums during the previous year and most of them have already visited the proposed museum (Palazzo Vecchio).

Table 4 reports by encouragement treatment, the average number of previous museums visit, the participation rate by class to visit the recommended museum and the average number of additional museums visit completed after 6 months : 40% of students who received the reward encouragement did visit Palazzo Vecchio, while only 10% of those who received the presentation and 3% of those who received the flier did complete the visit. The striking difference in the number of visits to the recommended museum among treatments is more evident if compared to the average number of visits completed during the previous year and during the following six months.

Table 4 – Further descriptive statistics across all students

	Average N. of Museum visit in previous year (stand dev)	*% of Visits to recommended museum*	*Average N. of Museum visit in the subsequent six months (stand dev)*
FLIER	3.27 (3.14)	3.23	1.49 (2.56)
FLIER & PRESENTATION	4.74 (4.03)	10.31	4.39 (4.22)
FLIER & PRESENTATION & REWARD	3.59 (4.15)	40.38	3.00 (2.50)
Total	3.87 (3.86)	18.71	2.95 (3.38)

3. A randomization inference approach for the analysis of a cluster-randomized experiment

In cluster-randomized experiments, randomization occurs at the cluster level; that is, the unit of assignment to an encouragement or a treatment is a group (e.g. Murray, 1998; Arcenaux, 2005; Raudenbush et al., 2007). In our study, the unit of assignment to the encouragements is the high school class. We assigned 15 different high school classes, each with an average of 19.8 students, to one of three treatments. Each treatment consisted of a different form of encouragement. All students in the same class received the same type of encouragement. Under these circumstances, the class is the natural unit of inference and standard methods for the analyses of randomized experiments can be

applied at the cluster level. By keeping the unit of analysis at the cluster level, we also avoid additional complications connected with individual analysis and stick to the goal of evaluating which form of encouragement works best, gross of the peer effect mechanisms that could be triggered by the encouragement itself.[5]

Let K=15 denote the clusters or classes, each containing n_k students. The classes were randomly and evenly assigned to the three forms of encouragement, so that, M_e= 5 (e = f, p, r) denote the five classes randomly assigned respectively to the flier, the presentation and the reward encouragement treatment. For each class k, let W_k denote the form of encouragement received: W_k= f for classes assigned to receiving the flier, W_k= p for classes assigned to receiving the flier and the presentation and W_k= r for classes receiving all three forms of encouragement. The variables of interest in the analysis are: 1) the share of students undertaking the encouraged visit to the proposed museums and 2) the class average number of subsequent visits to other museums undertaken in the next six months (with the encouraged visit excluded from this count). Both variables are defined at the class level. For each class, three potential outcomes can be defined for each variable of interest, however, only one outcome is observed associated with the encouragement level actually received.[6]

Indeed, the problem is that we can only observe one of the potential outcomes for each unit, depending on the encouragement level actually received. For example, if the class receives W_k= r, we can only observe the potential outcome $Y_k(r)$, and not $Y_k(f)$ and $Y_k(p)$, i.e. there is a problem of missing information about what the outcome of interest would have been had the class received an alternative encouragement. Our main interest is to compare the observable $Y_k(r)$ to the unobserved potential outcomes to establish whether, for class k and with respect to a given response variable, the reward encouragement works better than an alternative form of encouragement. Thus, the main challenge consists in finding the most credible approximation of $Y_k(f)$ or $Y_k(p)$. Once we find a way to achieve this goal, and if we assume that potential outcomes respond only to the encouragement and not to other features of the experiment (excludability assumption), then the effect on the kth class of receiving encouragement r rather than p or f can be approximated by the difference $Y_k(r)$ - $Y_k(e)$, with e=f or, alternatively, e=p. The credibility of the approximation depends on the process that determines which units

[5] It could be interesting to carry out the analysis at the individual level rather than the class level, by evaluating if and how the different forms of encouragements affect individual decisions. Such analysis would add the issue of intra-cluster correlation, which in our case corresponds to some sort of peer pressure that can affect both the individual decision to undertake the encouraged visit and the later voluntary museum attendance.

[6] We assume that there are no hidden variations in each level of encouragement and that the potential outcomes for any class do not vary with the encouragements assigned to any other class, as it would happen in the presence of between-cluster interference and externalities (*Stable Unit Treatment Value Assumption* SUTVA, Rubin, 1980).

receive which encouragement. Since encouragements are assigned at random by the experimenters, exogeneity is ensured, as assignments are independent of the unit's characteristics. The process of random assignment addresses the missing data problem by creating groups of observations that are, at least with regard to *a priori* expectations, identical. In this way, the missing potential outcomes for the *k*th class can be approximated with no systematic bias by the outcomes that are observed inother classes receiving an alternative level of encouragement (for a largely intuitive explanation, see Gerber and Green, 2012). Based on this, we can define the Average Treatment Effect (henceforth, ATE) of a particular form of encouragement over another as the average of all unit-level effects. For instance, the ATE for encouragement *r* with respect to encouragement *e=f* or, alternatively, *e=p* using a simple difference-in-means estimator (henceforth, DIM), is:

$$ATE_{r,i,DIM} = E[[Y_k(r)|\ W=r] - E[Y_k(e)|\ W=e]] = E[Y_k(r)|\ W=r] - E[Y_k(e)|\ W=e]. \quad [1]$$

With random assignment to clusters, it can be proven that the DIM is an unbiased estimator of the ATE (Imbens and Rubin, 2015; Gerber and Green, 2012). Randomization guarantees, with large samples, that pretreatment variables are well balanced in the subsamples defined by treatment assignments, and there is no need to adjust for background characteristics. However, with small samples, the process may lead to pretreatment variables not perfectly balanced. Since encouragement assignment is controlled by the experimenter and, therefore, exogenous, the background variables are not required for unbiased treatment effect estimation. However, it may happen that, although uncorrelated to the form of encouragement, background variables may affect the outcomes of interest, leading to somewhat imprecise estimates. To address the unbalances in background characteristics that persist after randomization, we chose to rescale the outcome by subtracting from the observed dependent variable its lagged value prior to (and thus independent of) the experiment, as the lagged value is very likely to be a good predictor of the outcome itself. We, therefore, used a difference-in-differences estimator (henceforth, DID) of the ATE. When comparing, for instance, *r* to *e=f* or, alternatively, *e=p:*

$$ATE_{r,i,\ DID} = E[[Y_k(r) - Y_{k,prior}]|\ W=r] - [Y_k(e) - Y_{k,prior}|\ W=e]. \quad [2]$$

The DIM estimator is also an unbiased estimator of the ATE and, since it has a smaller sampling variance than the DIM, it ensures precision gains (e.g. Gerber and Green, 2012).

We are now ready to outline the hypotheses from our randomization inference approach. The null hypothesis is:

H0: *The encouragement effect is zero for all classes, whatever encouragement form they receive, i.e.* $Y_k(f)=Y_k(p)=Y_k(r)$ *for all k.*

The alternative, one-tailed hypotheses are, respectively:

$H1_{p,f}$: *There exists at least one class k for which the "presentation" works better than the "flier" encouragement, i.e.* $Y_k(p) > Y_k(f)$ *or, equivalently,* $Y_k(p) - Y_k(f) > 0$.

$H1_{r,f}$: *There exists at least one class k for which the "reward" works better than the "flier" encouragement, i.e.* $Y_k(r) > Y_k(f)$ *or, equivalently,* $Y_k(r) - Y_k(f) > 0$.

$H1_{r,p}$: *There exists at least one class k for which the "reward" works better than the "presentation" encouragement, i.e.* $Y_k(r) > Y_k(p)$ *or, equivalently,* $Y_k(r) - Y_k(p) > 0$.

Let the ATE be our test statistic. We take the observed outcomes in our dataset, treat them as the missing potential outcomes and simulate all possible randomizations that could have taken place, so as to obtain an exact sampling distribution of the estimated ATE under the null hypothesis. In our sample, with five out of 15 classes assigned to each encouragement level, there are $[K! /(M_f! M_p! M_r!)] = [15! /(5! 5! 5!)] = 756{,}756$ possible randomizations. By looking at the distribution of these hypothetical ATEs, referred to as the randomization distribution, and contrasting it to the ATE estimated after the experiment, we can calculate the probability of obtaining – under *H0* – a fictional ATE that is at least as large as the one obtained from the actual experiment. This probability is equivalent to an exact *p*-value (Fisher, 1925) and represents a measure of extremeness. A small value (close to zero) of the *p*-value suggests that the observed value of the test statistic is very unlikely in the randomization distribution of the test statistic simulated under the null hypothesis of no effect. This constitutes evidence against *H0* and in favor of *H1*.

4. Results

Descriptive statistics at the class level from the field experiment, according to the form of encouragement received, are reported in Table 5: namely, the variables of interest (a and b) and an important background variable, the average (per class) number of museum visits undertaken during the twelve months preceding the experiment (c).
The table shows that the reward encouragement induces the largest share of students (40.2%) to visit the proposed museum, while participation rates are lower under the presentation (11.8%) and the flier (3.3%) encouragements. However, if we consider the

number of visit to museums in the six months following the experiment, then the number of average visits is higher in classes that received the presentation form of encouragement (4.303) than in those with the flier (1.756) or the reward (2.98).

Table 5 – Descriptive statistics at the class level by form of encouragement

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
FLIER					
a) share of students undertaking the encouraged visit	5	0.033	0.075	0.000	0.167
b) N. of voluntary museum visits 6 months later	5	1.756	2.013	0.000	4.875
c) N. of voluntary museum visits in previous 12 months	5	3.286	0.677	2.760	4.444
Difference (b - c)	5	-1.530	1.541	-3.222	0.431
PRESENTATION					
a) share of students undertaking the encouraged visit	5	0.118	0.263	0.000	0.588
b) N. of voluntary museum visits 6 months later	5	4.303	2.346	2.500	8.368
c) N. of voluntary museum visits in previous 12 months	5	4.749	1.237	3.760	6.857
Difference (b - c)	5	-0.446	1.303	-2.147	1.511
REWARD					
a) share of students undertaking the encouraged visit	5	0.402	0.180	0.222	0.650
b) N. of voluntary museum visits 6 months later	5	2.980	0.622	2.190	3.909
c) N. of voluntary museum visits in previous 12 months	5	3.543	1.058	2.667	5.000
Difference (b - c)	5	-0.564	0.679	-1.409	0.111

Table 5 also shows that, prior to the experiment, classes randomly assigned to the presentation encouragement reported a higher number of museum visits in the previous 12 months than the classes randomly assigned to the other two encouragement groups. This difference, due to chance, constitutes an example of unbalances in the background variables, despite randomization. Since the number of museum visits prior to the experiment is a reasonably good predictor of the number of museum visits during the experiment, it makes sense to consider the differences (b-c) and employ a DID in addition to a DIM estimator.[7]

Table 6 reports the results of the randomization inference analysis. Column (1) reports the ATEs observed values for each pair of encouragements; columns (2) shows the 95% confidence intervals calculated through the test-inversion procedure suggested by Rosenbaum (2002). The reward encouragement appears again as the most effective way to promote the one-time visit to the proposed museum, but not the best way to boost future museums' visits. The poor performance of the reward, relative to the presentation encouragement, comes out not only when using the DIM, but also with the more precise DID estimator.

[7]These differences have negative signs because the reference period prior to the experiment is 12 months, while the reference period during the experiment is six months. However, the related ATEs can be positive, since they are defined as the difference in Y between each pair of encouragement arms.

Table 6 – Observed encouragement effects and p-values for the null hypothesis H_0: $Y_k(1) - Y_k(0) = 0$ for all k and all pairs of encouragements

	(1) ATE	(2) ATE 95% C.I.		(3) p-value (right tail)
Effect on the share of students undertaking the encouraged visit (DIM)				
PRESENTATION vs FLIER	0.084	-0.023	0.208	0.293
REWARD vs FLIER	0.368	0.261	0.497	0.005
REWARD vs PRESENTATION	0.284	0.176	0.412	0.030
Effect on subsequent voluntary museum visits (DIM)				
PRESENTATION vs FLIER	2.547	1.494	3.810	0.016
REWARD vs FLIER	1.224	0.173	2.494	0.187
REWARD vs PRESENTATION	-1.323	-2.379	-0.053	0.836
Effect on subsequent voluntary museum visits (DID)				
PRESENTATION vs FLIER	1.084	0.031	2.347	0.015
REWARD vs FLIER	0.967	-0.084	2.236	0.189
REWARD vs PRESENTATION	-0.117	-1.173	1.152	0.840

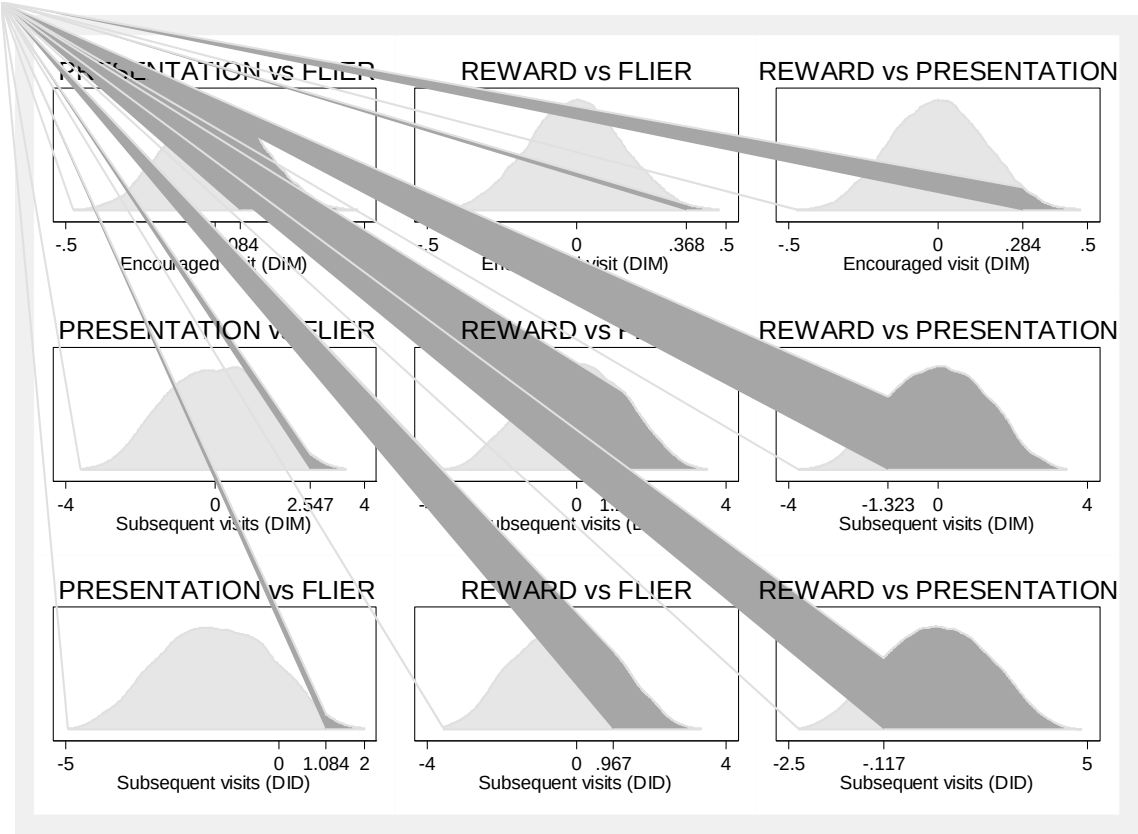
Column (3) shows the one-sided *p*-values, while the randomization distributions of the test statistic of interest, in the form of Kernel probability density functions, are shown in Figure 1.[8] For example, the observed differential effect (ATE) of receiving the reward rather than the presentation on the share of students undertaking the encouraged visit is 0.284, corresponding to 28.4% points higher under reward than under presentation. The probability to find, in the randomization distribution simulated under the sharp null hypothesis of no effect, a fictional value of this ATE that is at least as high as 0.284 is 0.030 (3%) (see column (3) in Table 6). In Figure 1, this probability corresponds to the dark grey area in the relative randomization distribution: the smaller the *p*-value (or, equivalently, the dark grey area), the stronger the evidence against the null hypothesis and in favor of the alternative one.[9]

Applying this argument to all encouragement pairs we can conclude that, to the end of promoting the one-time visit to the museum, the reward not only works better than the presentation encouragement, but also considerably better than the simple flier. On the other hand, there is no substantial difference between the effects of the presentation and the basic flier, as the probability of finding an ATE of 0.084 or higher in the randomization distribution is 29.3%.

[8]The smaller the dark grey region in the probability density function, the stronger is the support of the alternative hypothesis of some positive differential effect for at least one class.

[9]Note that a p-value of 0.030 also satisfies conventional requirements on statistical significance.

Figure 1 – The randomization distributions of encouragement effects for the sharp null hypothesis H_0: $Y_k(1) - Y_k(0) = 0$ for all *k* and all pairs of encouragements.



With regard to future visits, the observed ATE of receiving a presentation instead of a flier, using a DIM estimator, is 2.5, which means that the mean of future visits under the presentation is 2.5 points higher than under the flier encouragement. Since the related *p*-value is only 1.6%, we can infer that the presentation encouragement works. This differential effect narrows with the DID estimator, as we control for pre-experiment museum attendance: the difference in the number of visits after the presentation versus the flier is 1, but still highly significant (*p*-value = 1.5%). When contrasting the reward to the flier, the effect is still positive (the actual ATE = 0.9 with the DID estimator), but the probability of finding equal or greater values in the randomization distribution increases to 18.9%. Finally, we find no evidence of a differential effect of the reward versus the presentation to the end of boosting subsequent visits.

To sum up, the effect of the reward encouragement is large and significantly positive in affecting the one-time likelihood to visit to the proposed museum. The same is not true with respect to future visits. For this, we conclude that a policy based on extra credit rewards is not the most effective instrument to increase teens' attendance to museums over time and affect their long run behavior. Instead, classroom visits and direct

communication to students by museum educational staff appear to be a more successful means to increase and affect cultural consumption in teens and adolescents.

5. Concluding remarks

This paper reports results from a field experiment conducted to study incentives offered to high school teens to motivate them to visit art museums. This is a very novel area of application of field experiments. By means of a cluster-randomized trial, with randomization at the class level, we compared the immediate and the subsequent effects on museum attendance induced by three different types of encouragement: students in the low encouragement group received a flier containing basic information and opening hours of a main museum in Florence; students in the intermediate encouragement group receive the flier and a short presentation about the exhibit conducted by an art expert; students in the high encouragement arm, in addition to the flier and the presentation, receive also a non-financial reward in the form of extra-credit points towards their school grade. The analysis yielded two main findings. First, the extra-credit, non-financial reward, is more effective at inducing the students to undertake the encouraged visit than either the presentation or the basic information with the flier. Second, in a longer time horizon the extra-credit reward does not induce a significant change in behavior with respect to the simple presentation.

References

Arcenaux, K. (2005). Using cluster randomized field experiments to study voting behavior. *The Annals of the American Academy of Political and Social Science*, 601, 169–179.

Bakhshi, H., & Throsby, D. (2014).Digital complements or substitutes? A quasi-field experiment from the Royal National Theatre. *Journal of Cultural Economics*, *38*(1), 1-8.

Berlin, N., Bernard, A., &Fürst, G. (2014). Time spent on new songs: word-of-mouth and price effects on teenager consumption. *Journal of Cultural Economics*, *39*(2), 205-218.

Duflo, E., Glennerster, R., & Kremer, M. (2007).Using randomization in development economics research: A toolkit. In *Handbook of development economics*, edited by T. Paul Schultz and John A. Strauss, Vol. 4, 3895-3962.

Falk, M., & Katz-Gerro, T. (2015). Cultural participation in Europe: Can we identify common determinants?.*Journal of Cultural Economics*, to appear.

Fisher, R.A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, First edition.

Frangakis, C. E., Rubin, D. B., & Zhou, X. H. (2002). Clustered encouragement designs with individual noncompliance: Bayesian inference with randomization, and application to advance directive forms. *Biostatistics*, *3*(2), 147-164.

Gerber, A.S., & Green, D.P. (2012). *Field Experiments: Design, Analysis and Interpretation*. New York: W.W. Norton.

Gray, C. M. (1998). Hope for the future? Early exposure to the arts and adult visits to art museums. *Journal of Cultural Economics*, *22*(2-3), 87-98.

Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. New York: Cambridge University Press.

Levitt, S. D., & List, J. A. (2009). Field experiments in economics: the past, the present, and the future. *European Economic Review*, *53*(1), 1-18.

Mason, D. D., & McCarthy, C. (2006). 'The feeling of exclusion': Young peoples' perceptions of art galleries. *Museum Management and Curatorship*, 21(1), 20-31.

Murray, D. M. (1998). *Design and Analysis of Community Trials*. Oxford: Oxford Univ. Press.

Raudenbush, S.W., Martinez ,A., & Spybrook, J. (2007).Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, 29, 5–29.

Rosenbaum, P.R. (2002). *Observational Studies*. 2nd ed. New York: Springer.

Rubin, D. B. (1980). Discussion of Basu's "Randomization analysis of experimental data: The Fisher Randomization Test." *Journal of the American Statistical Association*, 75, 591-593.

Suárez-Vázquez, A. (2011). Critic power or star power? The influence of hallmarks of quality of motion pictures: an experimental approach. *Journal of Cultural Economics*, *35*(2), 119-135.