

Norm Enforcement in the city:  
A natural field experiment

Loukas Balafoutas and  
Nikos Nikiforakis

Nov 2011

Research Paper Number 1133

ISSN: 0819-2642

ISBN: 978 0 7340 4486 0

# Norm enforcement in the city: A natural field experiment<sup>\*</sup>

Loukas Balafoutas  
Department of Economics  
University of Innsbruck  
[loukas.balafoutas@uibk.ac.at](mailto:loukas.balafoutas@uibk.ac.at)

Nikos Nikiforakis<sup>†</sup>  
Department of Economics  
University of Melbourne  
[nikos.nikiforakis@unimelb.edu.au](mailto:nikos.nikiforakis@unimelb.edu.au)

November 22, 2011

## Abstract

The willingness of individuals to enforce social norms can have important welfare implications. All the experimental evidence on norm enforcement so far comes from laboratory experiments and indicates that many participants are indeed willing to enforce norms, even in one-shot interactions. We examine individuals' willingness to punish norm violations in a natural field experiment. We violate two efficiency-enhancing norms in the main subway station in Athens, Greece. The large number of passengers ensures that strategic motives for punishing are minimized. We find that some individuals punish norm violators, but the rate of enforcement is low relative to that found in laboratory experiments. Surprisingly, violations of the better known of the two norms are less likely to trigger punishment. Questionnaire data indicate that most people are concerned about being counter-punished, and that violators of the better known norm are considered more likely to counter-punish. Men are more likely to punish than women, while the rate of enforcement is unaffected by the violator's height and gender.

Keywords: norm enforcement, social norms, field experiment, altruistic punishment, cooperation  
JEL codes: C93, D63, H41

---

<sup>\*</sup> We would like to thank Klaus Abbink, Georgy Artemov, Marco Casari, Tim Cason, Glenn Dutcher, Andreas Lange, Simon Loertscher, Andreas Ortmann, Shyam Sunder, Matthias Sutter, Erte Xiao, as well as participants and seminar participants at the 2011 European Meetings of the Economic Society Association, the 6th Australian-New Zealand Workshop on Experimental Economics, the 10th TIBER Conference, the University of Melbourne and the University of Innsbruck. We would also like to thank Simon Gächter for sharing the data of Herrmann, Thöni and Gächter (2008). Funding from the Faculty of Business & Economics at the University of Melbourne is gratefully acknowledged.

<sup>†</sup> Corresponding author; telephone number: +61 383449717.

## 1. Introduction

Social norms can be defined as customary rules of behavior that govern interactions among people (Young, 2008). These rules influence our decisions in a wide range of circumstances ranging from the definition of property rights to our obligations towards our colleagues, family members and our behavior in public places. The interest of economists in social norms can be explained by the fact that, in many instances, norms serve to enhance efficiency by reducing externalities or transaction costs (e.g., Coleman, 1990; Elster, 1989; Young, 1998). Examples of efficiency-enhancing social norms include the norms of non-littering in public places and queuing.

The existence of social norms depends on the willingness of individuals to adhere to them. One reason why people may choose to adhere to a norm is because of the threat of being punished if they deviate from it (Sugden, 1986; Coleman, 1990).<sup>1</sup> Norm enforcement, however, is often costly for the enforcers as it can lead to reprisals. A typical assumption in economics is that, unless an individual stands to benefit directly from enforcing a norm and the benefit exceeds the cost, they will not do so. While punishing a norm violator can be beneficial if individuals interact repeatedly (e.g., Fudenberg and Maskin, 1986; Fudenberg et al. 1994), in large modern societies, many of the daily interactions are with strangers. This raises the question of what explains the adherence to norms that is often observed in one-shot interactions.

Recently, economists proposed an answer to this question: some individuals may be willing to punish norm violators even in one-shot interactions, because they derive non-pecuniary benefits from punishing (e.g. Fehr and Gächter, 2000; Fehr and Gächter, 2002; Masclet et al. 2003). In other words, some people may enjoy punishing norm violators so much that the benefit from punishment exceeds the cost. In anticipation of this, individuals may adhere to norms. The evidence in support of this explanation comes from laboratory experiments showing that many (indeed the majority of) participants are willing to pay to reduce the earnings of norm violators, even when they do not anticipate any pecuniary benefits themselves. For example, Fehr and Gächter (2002) report that 84.3 percent of subjects in their laboratory experiment use costly punishment in a public good experiment even though they know they will never interact again with the same individuals. As we will

---

<sup>1</sup> According to Young (2008), apart from the threat of punishment, there are two other mechanisms sustaining norms. One is negative emotions such as guilt or shame that are triggered when norms have been internalized. The other is the desire to avoid costs that could result from coordination failure.

see below, similarly high percentages are reported in most studies (for surveys see Chaudhuri, 2011; Gächter and Herrmann, 2009).

The prevalence of social norms in daily life implies that the willingness to engage in costly norm enforcement can have significant implications for a wide range of situations which are of interest to economists such as collective action and contract design (Fehr and Fischbacher, 2002). However, while there is considerable evidence from laboratory experiments regarding the willingness of individuals to engage in costly punishment, there is little evidence of such behavior in the field.<sup>2</sup> Indeed, anthropological studies indicate that costly punishment is rarely used for the enforcement of norms. Guala (2011; p.30) who reviews this literature concludes that “...there is no evidence in the anthropological literature that costly material punishment is used in small acephalous societies, except in the regulation of sexual conflict.” He suggests that the rare use of costly punishment in the field vis-à-vis the lab may be because important forces that exist in daily life (e.g., the ability of punished individuals to counter-punish or the ability to communicate with norm violators) are absent from most laboratory experiments.<sup>3</sup> However, a problem with the anthropological studies is that interactions are repeated and individuals have multiple ways for enforcing norms.<sup>4</sup> Thus, the rare occurrence of costly enforcement in these studies cannot provide conclusive evidence about whether individuals are willing to enforce norms at a cost in the field in one-shot interactions. Nevertheless, the absence of costly punishment in the anthropological literature raises the question of whether costly punishment is used by individuals to enforce social norms in one-shot interactions in daily life.

---

<sup>2</sup> Note that, while in many instances individuals can enforce norms through the use of reward or punishment, in this paper we focus on the use of costly punishment. As such, we will treat the terms *costly enforcement* and *costly punishment* as synonyms. Similar to Guala (2011), the adjective *costly* is used to indicate that the enforcer has to incur a cost for enforcing a norm and not that the cost of enforcing the norm exceeds the benefit.

<sup>3</sup> Allowing for these forces is known to limit the use of costly punishment (Denant-Boemont et al., 2007; Nikiforakis, 2008; Nikiforakis and Engelmann, 2011; Xiao and Houser, 2005). For example, similar to Fehr and Gächter (2002), Nikiforakis (2008) finds that when punishment cannot be retaliated and individuals are randomly reassigned to groups, 89.6 percent of subjects use costly punishment in a public good experiment. This percentage drops to 68.8 when punishment can trigger counter-punishment. Similarly, Nikiforakis and Engelmann (2011) report that when punishment cannot be retaliated, 81.3 percent of subjects punish, but when punishment can lead to a lengthy feud, 63.5 percent of subjects punish in the first punishment stage.

<sup>4</sup> Recent laboratory studies have provided evidence that individuals can be quite forward looking if there exists a prospect of future interactions (Cabral, Ozbay and Schotter, 2011; Reuben and Suetens, 2011). See also Casari (2011) for a discussion of the field evidence discussed in Guala (2011).

To address this question, in this paper, we present the results from a natural field experiment investigating whether individuals punish norm violators in anonymous, one-shot interactions. The experiment is “natural” in the sense that participants were unaware an experiment was being run (Harrison and List, 2004). To our knowledge, this is the first evidence from a natural field experiment on costly norm enforcement.<sup>5</sup> In the experiment, we violated two efficiency-enhancing social norms in a controlled way and recorded the behavior of those individuals who observed the violations. The “exogenous” violation of the social norms has the advantage that it allows us to control for the characteristics of the violators (e.g. gender, height, appearance) and is also necessitated by the fact that once a norm is well established violations are typically rare.

To ensure interactions are one-shot, the experiment was run in the main subway station in Athens, Greece. The station is used daily by hundreds of thousands of passengers which implies that strategic motives for punishing are minimized. The violations were done in a way that we have prior reason to believe makes punishment costly. Indeed, survey evidence that we collected and we discuss later in the paper indicates that punishment is considered costly by individuals. A further important feature of our design is that violations can only be punished by one person at a time. The reason we decided to include this feature in our study is that norm enforcement can be a second-order public good, since everyone benefits from the enforcement of the norm but all would prefer that someone else punishes violators, if punishment is costly.

Studying costly punishment in a natural field experiment has multiple advantages. First, participants in our study do not self select into the experiment. We thus avoid a potential selection bias which can limit the external validity of experimental results (e.g. Harrison and List, 2004; Levitt and List, 2007). Second, participants are neither aware that their actions are being observed nor do they have beliefs about the aim of the experiment. As is well known from economic and psychology experiments, both factors can significantly affect participants’ behavior (e.g., Benson, 2000; Hoffman, McCabe and Smith, 1996). Third, the cost of punishment in the field is arguably higher. If a violator decides to avenge punishment in the lab, at most, the enforcer can lose their entire earnings from the experiment. In contrast, in the field, the consequences can be dire; in extreme cases, punishing a violator can lead to the

---

<sup>5</sup> In a recent paper, Noussair, van Soest and Stoop (2011) report the results from a framed field experiment to examine cooperation in a social dilemma where individuals can use costly reward and punishment.

enforcer being injured, psychologically traumatized or even killed.<sup>6</sup> One disadvantage of field experiments such as ours is that some control is inevitably lost; for example, we cannot control the relative cost of punishment for each individual. However, this can also be seen as an advantage as we observe individuals' willingness to punish in a setting where the cost of punishment is not induced by the researcher. In addition, as we discuss in more detail below, to gain insights into participants' motivation, we also conducted surveys asking them about the reason for their actions.

Both norms that we violated for the purposes of the experiment are well established, but differ in the degree of their universality. The first norm prescribes that passengers stand on the right side of the escalators so that others who wish to walk up the escalators can use the left side. One of the experimenters or the experimenters' associates stood on left side. In case an individual requested the norm violator to stand on the other side, the violator was instructed to ignore the request. This implies that enforcers would have to insist and possibly raise their voice or push the violator aside, thus making norm enforcement costly for them. This norm is not universal but *environment specific* as, in Greece, it is found only in the subway system. The second norm, on the other hand, is *universal*. It prescribes that individuals do not litter in public places. The non-littering norm is a *universal* as it exists not only in subway stations, but also in other places (e.g., at home, at the workplace). All subway stations in Athens are noticeably litter-free. Violators of the non-littering norm threw either a wrapped-up colored piece of paper or a plastic water bottle in a corridor leading to the platforms. Our hypothesis is that the universal norm will be better known to individuals and thus violators of it will be more likely to be punished. The rationale behind this hypothesis is that violations of the universal norm are more likely to be viewed as intentional. Note that both norms are efficiency enhancing as most passengers benefit from a clean subway station and would prefer to have the option to walk on the left side in case they are in a hurry.

Our findings indicate that, in line with many laboratory experiments, some people are willing to use costly punishment to enforce norms. However, the rate of enforcement is low relative to the number of individuals punishing norm violators in the laboratory. Out of 300 cases of norm violation, punishment was observed in only 35 cases (11.7 percent).

---

<sup>6</sup> Recently there have been a number of high-profile cases in Germany and the U.K where individuals were either severely beaten or killed for trying to enforce social norms. In one case in Germany, for example, a pensioner was severely beaten for asking two teenagers to stop smoking in a Munich subway station (Spiegel, 2008). In another case in England, police officers were beaten up after asking a 15-year old girl to stop littering (Edwards, 2011). For other recent examples see Fresco (2008) and Spiegel (2010).

Questionnaire data indicate that the main reason for the low rate of norm enforcement is that people are concerned about being counter-punished by the norm violator. The rate of enforcement is affected by neither the height nor the gender of the violator, but men are more likely than women to enforce norms. Perhaps our most surprising finding regards the different rate at which the two norms are enforced. In contrast to our expectations, violators of the universal norm are substantially less likely to be punished than those of the environment-specific norm (4 percent versus 19.3 percent). Our questionnaire data suggest that the reason for this is that individuals who litter are considered as being more likely to counter-punish than those who stand on the wrong side of the escalators. We conjecture that this is because violators of a widely shared norm are more likely to be generally antisocial and, hence, more likely to react negatively to an individual's attempt to enforce upon them a social norm.

The paper proceeds as follows. In the next section, we discuss the experiment in detail. In section 3, we present the results from the experiment. The last section concludes by discussing the implications of our findings and topics for future research.

## **2. The experiment**

### **2.1. Location and population**

The experiment was run in the main subway station of Athens (Syntagma Station). Approximately 650,000 passengers use the two main subway lines of the Athenian subway which intersect at the station.<sup>7</sup> As described below, the sessions were run in five different locations inside the station. The team of experimenters consisted of six individuals (the researchers and four research assistants), whose tasks included the systematic violation of two social norms, the collection of data on the behavior of passengers in response to violations, as well as the collection of survey data inside the station. All sessions were run between 2pm and 8pm (in order to avoid morning peak hours) on four different working days between April 27<sup>th</sup> and May 3<sup>rd</sup>, 2011.

One may wonder how different the population in Athens is compared to other “populations” frequently studied by experimental economists (e.g., British, Swiss) with regards to its inclination to enforce norms. To address this question, prior to running the experiment, we used the data collected by Herrmann, Thöni and Gächter (2008) to compare the rate of costly punishment in laboratory experiments in Athens vis-à-vis other locations. The results can be found in Table A1 where we present the percentage of subjects that engage

---

<sup>7</sup> The information was taken from the ATTIKO METRO website on July 23, 2011.

<http://www.amel.gr/index.php?id=8&L=1&cHash=37a860319ea26b2e30e3b58f6b6ecfdc>

in costly punishment (our variable of interest) and the percentage of instances where punishment was directed towards an individual contributing less than his/her peers on average. It can be easily seen that subjects in Athens are not noticeably different than subjects in other places. For example, consider Nottingham and Zurich – two laboratories where experiments on costly punishment have frequently been run. We find that the percentage of subjects using costly punishment in laboratory experiments in Athens (88.6 percent) is similar to that in Nottingham (89.5 percent) and Zürich (82.6 percent). The majority of punishment cases in Athens (62.8 percent) is aimed towards free riders who contribute less than their peers on average. This is similar to the percentage of punishment cases aimed at free riders in Nottingham (61.7 percent) and slightly lower than in Zürich (69.9 percent). On the basis of this evidence, we have no prior reason to believe that the rate of enforcement in Athens should be noticeably different than in other locations that have been frequently studied by experimental economists.

## **2.2. Social norms and violations**

### **2.2.1. Escalators**

The “escalators norm” prescribes that passengers stand on the right side of the escalators so that others who wish to walk up the escalators can use the left side. Violations of the escalators norm were implemented as follows. An experimenter (henceforth, the *violator*) was waiting on the platform once a train arrived. Violators were dressed casually wearing a pair of new blue jeans and a black shirt. After a train arrived and the first transit passengers stepped on the escalator, the violator started walking upwards, on the left side, as the norm dictates.<sup>8</sup>

After taking a couple of steps, the violator stopped walking and stood on the left side of the escalator next to another passenger on the right side. This was done in a way that made it difficult for those behind the violator to continue walking upwards. After waiting for approximately 8 seconds, the violator moved to the right side of the escalators, allowing the passenger(s) behind them to pass. If before the 8 seconds elapsed the individual immediately behind the violator (henceforth, the *observer*) asked the violator to move to the right, the violator was instructed to ignore the request and keep standing on the left. This was done to

---

<sup>8</sup> The number of passengers who stepped on the escalator before the experimenter was between 15 and 20. This was done for three reasons. First, it allows transit passengers to observe the norm before the violation occurs. Second, it reduces the likelihood that the person behind the violator is in a hurry, as these individuals are typically the first to come out of the trains – an issue which we discuss in greater detail later in this section. Finally, after the first passengers reach the escalator, a longer queue is quickly formed on the right side which implies that individuals are not indifferent between standing on the left and the right side, all else equal.



ensure that enforcement was considered to be costly for observers. If the observer asked for a second time, then the violator moved to the right side of the escalator without saying anything. In those cases where the observer insisted on the violator standing to the right, we say that they enforced the norm. We call these observers *norm enforcers* or simply *enforcers*.<sup>9</sup>

We collected data on three different escalators, all of them connecting the two main lines of the subway. The escalators were all moving upwards, taking passengers from their arrival to their departure platform. Next to each escalator there was also a staircase, which could be used by passengers who were in a hurry to catch a train. As we are interested in studying the propensity to engage in costly norm enforcement, the experiment was designed in a way that minimized the number of observers who would punish because they were in a hurry. In particular, five steps were taken. First, the experiment was run in relatively short escalators where the private benefit from forcing a violator to stand on the right was minimal (about 8 seconds). Second, the experiment was run during off-peak hours so that fewer people were in a hurry. Third, during these hours, trains ran every three minutes, so that missing a train did not impose a large waiting cost on observers. Fourth, we only chose escalators with a staircase next to them, which could be used by passengers who were in a hurry to catch a train. Finally, experimenters waited for a while before they violated the norm, so that passengers in a hurry could rush to the escalators. We return to this topic in the results section, where we show that the large majority of instances where an observer asked the violator to stand on the right were indeed due to a desire to enforce the escalators norm, and not due to observers being in a hurry.

Experimenters worked in pairs. Besides the violator, a second experimenter stood discretely at the top of the escalator and collected data regarding: the location where the violation occurred; the time of the violation; the identity of the violator; the sex, approximate age, height and the reaction of the observer. The latter included information on whether there was norm enforcement, whether the observer continued walking up the escalator when the violator stepped to the right, and – if there was no enforcement – whether the observer tried to pass without talking to the violator.

---

<sup>9</sup> While in principle other passengers could also enforce the norm by shouting at the violator from a distance, it seems natural that the responsibility for doing so rests with the person directly behind the violator, whose way is obstructed. This is actually one of the reasons why we chose to violate this norm, as the second-order public-good problem is minimized. In any case, we almost never observed any cases of enforcement from passengers other than the enforcer.

In order to have a better understanding of the motivation of enforcers, once they reached the top of the escalators, the second experimenter approached them and asked them a small number of questions (which in many cases meant that they had to board the train along with the enforcer). In particular, the enforcer was asked to indicate the reason for asking the violator to stand aside. This could be (i) that they were in a hurry; (ii) that it is not right to stand on the left side of an escalator; or (iii) that they wanted to teach the violator that in the future they should stand on the right. Options (ii) and (iii) are meant to help us distinguish between those who are simply annoyed by the violation of the norm and those who are more forward looking. Multiple answers were allowed. Enforcers were also allowed to express other reasons of their own for asking the violator to stand aside.

### 2.2.2. Littering

The second norm prescribes that individuals do not litter in public places. We violated the non-littering norm inside the subway station by throwing either a wrapped-up colored piece of paper (A4 size) or a small, empty plastic bottle of water (approximately 30 cm long) in a corridor leading to one of the platforms. The violator pretended to be reading a sign explaining the location of the different platforms, until they noticed a *single* person approaching. That person will be called the *observer* in the littering treatment, and we will say that an observer enforced the norm if they asked the violator to pick up the garbage they threw or reprimanded them in some other way. That is, as in the case of the escalators, norm enforcement is not a second-order public good.<sup>10</sup>

Once the observer approached the violator, the latter started walking in front of them at a pace that was slightly slower than the pace of the observer. When the observer was approximately 3-4 meters from the violator, the latter threw the paper/bottle on the ground. The violation was done in a way that left no doubt that the violator intended to litter. However, the fact that the observer was behind the violator meant that the observer could not be sure whether the violator knew that they were being observed.<sup>11</sup> We decided to do this as we thought that the non-littering norm would be less likely to be enforced if it was violated in a provocative manner.

---

<sup>10</sup> As in the case of the escalators, other individuals were present in the area where the violation occurred. This implies that an observer could reasonably expect others to interfere if tension was to be triggered by enforcing the norm. Of course, this was not possible in our experiment, but observers were not aware of this.

<sup>11</sup> Videos of the two types of violations are available from the authors upon request. Other material related to the experiment such as the questionnaires used can be downloaded at <http://www.economics.unimelb.edu.au/staff/nikosn/>.

Experimenters again worked in teams of two. The second experimenter stood approximately 10 meters away from the point of the violation pretending to read. Apart from recording information regarding the time and the location of the violation, as well as the characteristics of the observer (see section 2.2.1), their role was to ascertain that the observer witnessed the norm violation by monitoring their gaze. In case an observer enforced a norm, the second experimenter went up to them and asked them to respond to a short questionnaire examining the reason(s) for enforcing the norm. Possible answers included a desire to indicate to the violator that they should respect the clean environment in the future, that it's wrong to throw litter, or some other reason.

As mentioned, our main motivation for choosing the two particular social norms was that they vary in the degree to which they have a universal status within the society. While the days we spent planning and running the experiment convinced us that both norms are well-established in the sense that they are adhered to by the vast majority of people, we expected that the non-littering norm would be better known as it is universal. The survey responses confirm this intuition. When asked about the share of passengers who responders believe are aware of the non-littering norm inside the subway, 89.3 percent of respondents indicated their *belief* that most or all passengers were aware of the norm. This percentage is markedly lower for the environment-specific escalators norm (33.3 percent). Note that the relatively low percentage should not be interpreted as evidence that the escalators norm does not exist. In fact, we show later that the majority of responders are annoyed by the violation of the escalators norm. Hence, we believe we have been successful in selecting two norms with a clear difference in their degree of universality.

### **2.3. Additional treatments**

Apart from examining the impact of the universality of the norm, we studied the impact on norm enforcement of the height of male violators in both the escalators and the littering treatments. Male violators were either 1.70m or 1.90m. At the time of the experiment, the average height for a Greek male was 1.80m. It seems plausible that punishing a tall violator will be perceived by enforcers as being more costly all else equal, given the risk of direct confrontation in the event of counter-punishment by the violator.

We also varied the gender of the violator. The female violator was 1.70m.<sup>12</sup> Interpreting differences in enforcement rates depending on the gender of the violator is not

---

<sup>12</sup> We did not vary the height of the female violator as we were unable to find female research assistants of 1.50m. The average height for a Greek female is 1.60m.

straightforward. Higher enforcement rates could be due to the fact that enforcers perceive the punishment of female violators as being less costly. However, other gender-related factors such as a reluctance to punish a female violator due to chivalry may reduce enforcement rates.

The resulting 2x3 design, summarized in Table 1, consists of two different social norms and three different types of norm violators. We collected 50 observations per cell, summing up to a total of 300 observations.

**Table 1: Experimental Treatments**

<i># of obs.</i>	Male violator: 190 cm	Male violator: 170 cm	Female violator 170 cm	Total
Escalators	50	50	50	150
Littering	50	50	50	150
Total	100	100	100	300

#### 2.4. Surveys

In order to better understand behavior in the experiment, we also conducted two surveys on general attitudes towards the two social norms in question. The surveys were conducted in the same subway station where the experiment took place.

Survey respondents were initially faced with a scenario describing the violation of either the escalators norm or the non-littering norm. The scenario mirrored the way in which norms were violated in the experiment using, however, a neutral language (i.e., no reference was made to ‘norms’ or ‘violations’). Respondents were then asked if they would be bothered by the actions described in the scenario. In case of a positive response, respondents were further asked if they would confront the individual in the scenario. Those who stated that they would not be willing to do so were then asked to indicate their reason(s), which could include the fear that norm enforcement could lead to tension (i.e., counter-punishment), the fact that the norm is not widely accepted, the fact that others do not enforce the norms, and other norm-specific questions.<sup>13</sup> The order with which the different reasons appeared in the survey was randomized. Also, respondents were invited to offer different reasons for enforcing the norm.

<sup>13</sup> For example, in the case of littering, one response was that there is cleaning staff employed by the company in charge of the subway. In the case of the escalators, respondents could state that the delay imposed by the violator is trivial.

In addition, the surveys elicited data on the perceived universality of each of the two norms, the frequency of using the subway in Athens and other cities, and some demographic data. In total, we conducted 150 surveys; 75 for each of the norms. 50.7 percent of the respondents were female. The mean estimated age among respondents was 37 years, while the median age in the population is 42.5 years. Finally, as already explained in section 2.2.1, we also conducted a survey targeting observers who enforced the escalators norm.

### 3. Results

The rate of norm enforcement is positive, but considerably lower than that observed in laboratory experiments. Out of the 300 cases when one of the norms was violated, there were only 35 instances of enforcement. That is, norms were enforced in 11.7 percent of instances. Figure 1 shows that, contrary to our expectations, the universal norm (non-littering) is substantially less likely to be enforced than the environment-specific norm (escalators). In particular, the non-littering norm was enforced in merely 6 out of 150 cases (4 percent). In contrast, the escalators norm was enforced in 29 out of 150 cases (19.3 percent).<sup>14</sup> The difference in enforcement rates across the two norms is statistically significant ( $p$ -value $<0.01$ , chi-square test).

**Result 1:** *Norm enforcement is observed in 11.7 percent of cases. The universal norm is nearly 5 times less likely to be enforced than the environment-specific norm.*

What can explain the relatively low rate of enforcement and the relative infrequency with which the universal norm is enforced? To answer these questions we turn to our evidence from the surveys. The first explanation one needs to consider is that individuals are simply not sufficiently bothered by the violations of the norms and that they are less bothered by violations of the non-littering norm than the escalators norm. As it turns out, this explanation cannot account for the findings summarized in Result 1.

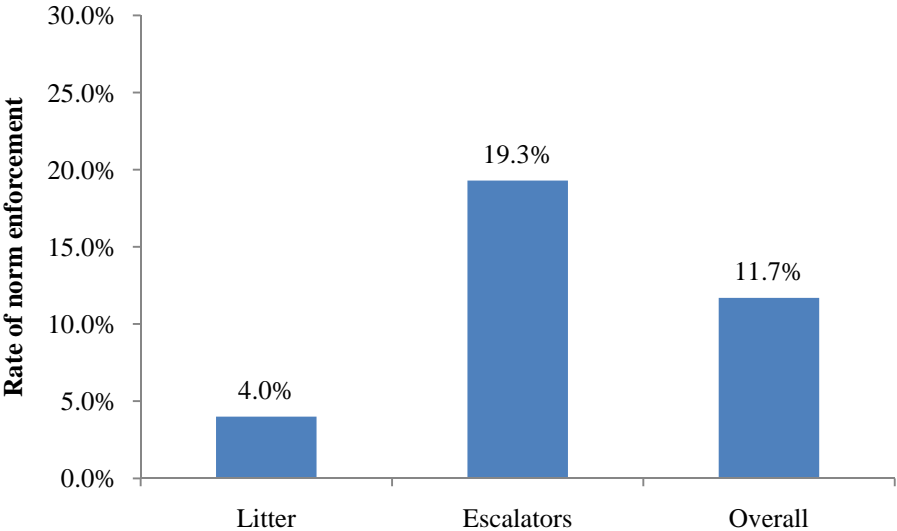
The majority of respondents stated that they would be bothered “quite a lot” or “a lot” by violations of the norms (68 percent – 102 out of 150 respondents). In particular, 45.5 percent of the respondents said they would be bothered by an individual standing on the left side of the escalators (34 out of 75 respondents), while 90.7 percent of survey respondents stated that

---

<sup>14</sup> There were also 8 instances in which the observer asked the violator to move to the right, but did not insist after the violator ignored their request. We do not classify these instances as costly norm enforcement. However, none of the results reported in the paper are affected if we add these observations in the analysis.

they would be bothered by a person who litters in the subway (68 out of 75 respondents).<sup>15</sup> The difference is statistically significant ( $p$ -value<0.01, chi-squared test). Therefore, a paradox emerges where violations of the universal norm trigger stronger negative reactions – as one would have anticipated – but are *less* likely to trigger punishment.<sup>16</sup>

**Figure 1: Rates of norm enforcement in the experiment**



The most common reason given by respondents for not enforcing the norm when a violation is observed – despite having stated that the violation would bother them – is their fear of retaliation by violators (57.8 percent of respondents across norms). This indicates that punishment is considered by individuals to be costly. Interestingly, the percentage of individuals who fear that punishing a norm violator could trigger counter-punishment is substantially larger in the case where the violator litters than in the case where they stand on the left side of the escalators. This can be seen in Figure 2 that presents the percentage of individuals who are bothered by the violation of the norm but are unwilling to punish the norm violator out of fear of counter-punishment. In particular, 68.3 percent of respondents stated that they would fear being counter-punished by an individual who litters. In contrast,

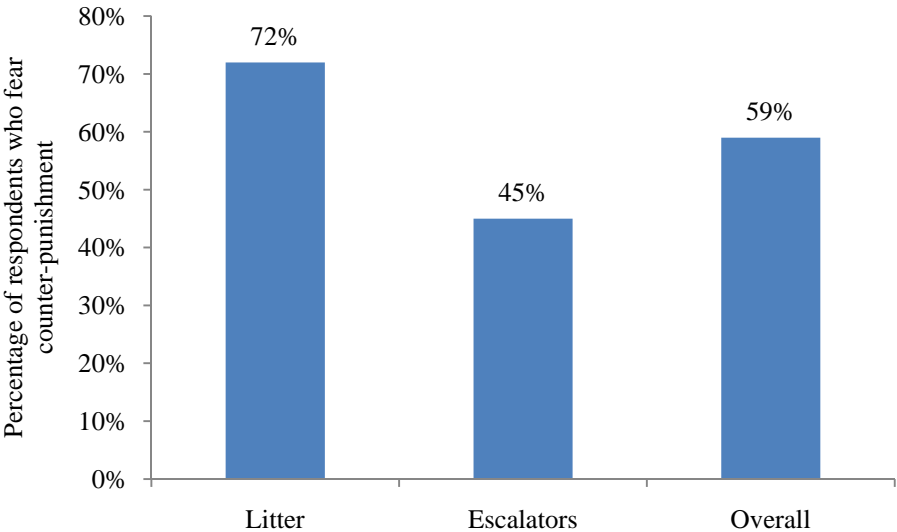
<sup>15</sup> This evidence seems to also refute another possible explanation for the lower rate of enforcement across the two norms, namely, that enforcement of the non-littering norm is more like third-party punishment in laboratory experiments and enforcement of the escalators norm more like second-party punishment. If this was the case, based on laboratory evidence (Fehr and Fischbacher, 2004), one would expect that observers of violations of the non-littering norm would be *less* bothered than observers of the violations of the escalators norm.

<sup>16</sup> It is worth noting that, while survey respondents are more likely to state that they would punish norm violations than observers did in the experiment, in line with our experimental results, survey respondents are found to be more willing to enforce the escalators norm (29.3 percent) than the non-littering norm (18.7 percent). Therefore, the survey responses yield qualitatively the same result as our experiment.

42.9 percent of respondents stated that they would fear counter-punishment by an individual who stands on the left side of the escalators. The difference is statistically significant ( $p$ -value=0.01, chi-square test).

**Result 2:** *The main reason for the relatively low rate of enforcement in the experiment is the fear of retaliation by violators. Violators of the universal norm (non-littering) are deemed to be more likely to retaliate than violators of the environment-specific norm (escalators).*

**Figure 2: Percentage of respondents who would not punish norm violators out of fear of counter-punishment.**



Why would violators of the universal norm be considered as more likely to retaliate punishment? One explanation can be that norm violations reveal information about the type of individual violating the norm. It seems plausible that the more well-known a norm is (as is the case with the universal norm; see survey evidence above) the more antisocial the individual violating the norm is considered to be. Observers of such violations may reasonably infer that the violator generally disregards social conventions, so that an effort by an observer to enforce the norm is more likely to trigger counter-punishment. Nikiforakis and Engelmann (2011) provide evidence from a laboratory experiment in line with this explanation. They find that, when punishment in a public good experiment can lead to a long sequence of punishment and counter-punishment, extreme free riders are *less* likely to be punished than less extreme free riders. Their evidence further indicates that extreme free riders were indeed more likely to respond to punishment by counter-punishing in their experiment. In our case, littering is an apparent violation of a widely-known social norm. In contrast, the escalators norm is environment specific. An individual that is an infrequent user of the subway may not be aware

of the social norm. If someone who litters is perceived as more likely to counter-punish as indicated by responses in the survey, then the expected cost of enforcing the norm is higher and, hence, the rate of norm enforcement should be lower, all else equal (Anderson and Putterman, 2006; Carpenter, 2007; Egas and Riedl, 2007; Nikiforakis and Normann, 2008). This can explain why norm enforcement is lower in the case of the non-littering norm – the universal norm.

**Conjecture:** *The lower rate of enforcement in the case of the universal norm is due to the perception of observers that violators of universal norms are more likely to be antisocial in general and thus more likely to respond to punishment by counter-punishing.*

Another plausible explanation for the higher rate of enforcement in the case of the escalators norm is that the direct benefit from enforcing the escalators norm exceeds that from enforcing the non-littering norm. This seems plausible as violators of the escalators norm may cause some observers to miss the next train. As discussed in section 2.2.1, a number of steps were taken to minimize the number of observers who were in a hurry. Nevertheless, it is still possible that some of the “enforcers” in the case of the escalators were simply in a hurry. In order to address this concern, we examine the reasons provided by those enforcing the escalators norm for their actions in the separate survey that we conducted with them (see section 2.2.1).

Of the 29 enforcers, we were able to obtain responses from 23; three enforcers refused to respond to the survey without stating a reason, one enforcer did not understand either Greek or English, and two enforcers boarded the train before the experimenter could reach them. The majority of enforcers were driven by a desire to enforce the escalators norm (74 percent – 17 out of 23 enforcers). Only 5 of the respondents stated that they were in a hurry as the sole reason for asking the violator to stand to the right (22 percent). If we ignore the six non-respondents, this evidence implies that 74 percent of the instances where the escalators norm was enforced reflected costly norm enforcement. We could therefore say that costly enforcement occurred in 14.3 percent of cases (19.3 percent x 74 percent). This rate is still more than three times higher than that in the case of the non-littering norm (4 percent). We therefore conclude that the higher rate of enforcement of the environment-specific norm is not a result of observers being in a hurry.

Before we proceed to examine other factors that impact on the willingness of individuals to engage in costly norm enforcement, for completeness, we briefly discuss some of the other reasons given by survey respondents for not punishing norm violators. In the case of the



escalators norm, 13 percent of respondents indicated that they would not enforce the norm despite being bothered as the delay is short and, therefore, the benefit from enforcing the norm is low. Also, 12 percent of respondents said that there is no clear norm prescribing that individuals stand on the right side of escalators. In the case of littering, the second most common reason given for the unwillingness to enforce the non-littering norm (after the fear of counter-punishment) is the existence of staff whose job is to keep the subway clean. This response was given by 7 percent of respondents.

We now turn our attention to the impact that the height and gender of the violator have on the observers' willingness to enforce norms. The variation in height is interesting as one could expect that, the taller a violator, the more costly norm enforcement will be, all else equal. However, as it turns out, the height of the violator does not impact on the rate of norm enforcement. In particular, taller male violators (1.90m) are punished in 10 out of 100 cases, while shorter male violators (1.70m) are punished in 12 out of 100 cases. The difference is not statistically significant ( $p$ -value=0.65, chi-squared test).

**Result 3:** *The height of the violator does not affect the likelihood of rate enforcement. Taller violators (1.90m) are as likely to be punished as shorter violators (1.70m).*

Since the fear of counter-punishment is a major determinant of whether to punish violators, one may expect that observers will be more likely to enforce a norm when the violator is a woman since they may be less likely to counter-punish and less effective in doing so. However, at the same time, observers may be unwilling to punish women as, similar to many other developed countries, a social norm of chivalry exists in Greece prescribing that women are treated with more care than men. As it turns out, the gender of the violator does not affect the likelihood that a norm is enforced. In particular, controlling for the height of the violator (1.70m), female violators are punished in 13 out of 100 cases, while male violators in 12 out of 100 cases. The difference is not statistically significant ( $p$ -value=0.83, chi-squared test).

**Result 4:** *The gender of the violator does not affect the likelihood of norm enforcement. Female violators are as likely to be punished as male violators.*

The majority of enforcers in our experiment are male (23 of 35 across norms). Given that the sample of observers is quite gender-balanced (152 women and 148 men), the gender difference in punishment rates among enforcers is statistically significant ( $p$ -value = 0.04, chi-square test). This result is mainly driven by enforcers in the escalators treatment as there are

very few instances in which the non-littering norm was enforced. It is interesting to see whether male respondents in our survey are also more likely to state that they would punish norm violations than female respondents. This is indeed the case. Across norms, 23 out of 74 male respondents said they would enforce the norm (31.1 percent), compared to 13 out of 76 female respondents (17.1 percent). The difference between these “hypothetical” enforcement rates is statistically significant ( $p$ -value=0.05, chi-squared test). A likely reason for the difference in the propensity to enforce norms between men and women seems to be the fear of counter-punishment: female respondents gave the fear of counter-punishment as a reason for not enforcing a norm in 63.2 percent of the cases where they would be bothered by the violation of the norm but would not enforce it (36 out of 57 cases). In contrast, 51.1 percent of male respondents gave the fear of counter-punishment as a reason for not enforcing a norm. However, this difference is not statistically significant ( $p$ -value=0.22, chi-squared test).

**Result 5:** *Men are more likely to punish norm violators than women.*

Before we conclude this section, we present in Table 2 evidence from a regression analysis investigating the determinants of norm enforcement in our experiment. The dependent variable is a binary variable taking the value of 1 if the observer enforced the norm and 0 otherwise. The first regression includes our treatment variables (non-littering norm, violator’s height and violator’s gender) as independent variables. The second regression extends the empirical model to control for the observable characteristics of the observer, namely whether they were male or female, and whether they were shorter than the violator. Entries are marginal effects.

The results presented in column (1) confirm Results 3 and 4, that is, the height and the gender of the violator do not affect significantly the likelihood that the observer punishes the violator. In addition, we find that violators of the non-littering norm are approximately 15 percent less likely to be punished. The results presented in column (2) provide additional support for Result 5. In particular, male observers are 9.3 percent more likely to enforce a norm. The regression also shows that the height of the observer relative to the violator does not impact on the likelihood of norm enforcement.

**Table 2: Determinants of norm enforcement**

<i>Dependent variable:</i> Probability of norm enforcement	(1)	(2)
<i>non-littering norm</i>	-0.153 *** (0.036)	-0.150 *** (0.036)
<i>male violator</i>	-0.030 (0.045)	-0.058 (0.053)
<i>tall violator (190cm)</i>	-0.016 (0.043)	-0.032 (0.049)
<i>male observer</i>		0.093 ** (0.038)
<i>observer shorter than violator</i>		0.035 (0.042)
$R^2$	0.088	0.115
<i>prob. &gt; <math>\chi^2</math></i>	0.001	0.000

Results are from a probit regression with robust standard errors; entries are marginal effects. N=300; \*\*, \*\*\* denotes significance at the 5%, 1% level respectively.

#### 4. Discussion

The willingness of individuals to adhere to social norms in one-shot interactions and generally cooperate with strangers has been a long-standing puzzle for economists and other social scientists. An explanation that has received considerable attention recently is the willingness of individuals to engage in costly enforcement even when they cannot anticipate any benefits from their actions. The supporting evidence comes from laboratory experiments demonstrating that most participants are willing to sacrifice part of their earnings to reduce those of others in one-shot interactions. However, costly punishment appears to be extremely rare in observational field studies, thus raising the question of whether costly punishment occurs and can sustain social norms in one-shot interactions (Guala, 2011).

We have presented results from a natural field experiment designed to examine whether individuals are willing to enforce social norms at a personal cost in one-shot interactions. The experiment was run in the main subway station in Athens, Greece, violating two efficiency-enhancing norms in a controlled way. Participants were unaware an experiment was being conducted and the large number of passengers travelling through the station implies that strategic incentives for enforcing the norm were minimized.

Consistent with laboratory evidence, we have found that some individuals are indeed willing to enforce norms at a cost in one-shot interactions. However, the rate of enforcement in our field experiment is considerably lower than that typically observed in the laboratory. The main reason for the lower rate of enforcement in our field experiment appears to be the fear of retaliation by the violator. Our study also reveals some interesting and surprising facts about the determinants of norm enforcement in the field. In particular, contrary to our expectations, the better known of the two norms (non-littering) is less likely to be enforced, despite the fact that individuals appear to be more annoyed by the violations. Our survey evidence suggests that the reason for this is that violators of well-known norms are considered to be more likely to avenge punishment. It seems intuitive that individuals who disregard conventions that are known to be widely shared by others will be more likely to react negatively to any attempts of enforcement. In contrast, some individuals who violate less well-known norms (such as the escalators norm in our experiment) may do so because they are oblivious of the norm and, hence, less likely to avenge punishment. Of course, we cannot rule out the possibility that other factors may partly explain the difference in enforcement rates across the two norms. While we have taken steps to make enforcement conditions comparable across the two norms, it is possible that some unobservable factors differ across the two norms and can partly account for the lower rate of enforcement of the non-littering norm. Finally, we have also found that the height and the gender of the violator do not seem to affect the rate of enforcement, while men are more likely to punish violators.

The rate of enforcement in our field experiment – which is higher than that reported in observational studies, but lower than that in laboratory experiments – raises two questions for future research. The first question is what can explain adherence to social norms in one-shot interactions. Our experiment was not designed to address this question, since, in order to control for some of the determinants of enforcement, violations were exogenous. However, it is worth mentioning that the norms studied appear to be widely adhered to by the population investigated: in the four days during which the experiment was run, we witnessed no violation of the non-littering norm (or any evidence of littering), and very few violations of the escalators norm.<sup>17</sup> Given the positive rate of enforcement, it is possible that part of the adherence can be attributed to the fear of punishment. This rate may be sufficient (indeed, optimal) to maintain norms once they have been firmly established. However, other factors

---

<sup>17</sup> It is also worth mentioning that 90.9 percent of the observers who did not enforce the norm proceeded to walk up the escalators once the violator stepped to the right. This shows that individuals adhered to the escalators norm even after they observed others violating it.

are likely to play a role too (e.g., conformity, internalization of norms), as laboratory experiments have shown that even substantially higher rates of costly punishment cannot lead to full cooperation such as that observed in the case of the non-littering norm. For example, Nikiforakis and Engelmann (2011) find intermediate levels of cooperation in a public good experiment even though 63.5 percent of subjects use costly punishment after observing individual contributions towards a public good. Also, one cannot rule out the possibility that costly punishment plays a greater role during the period that a norm is established and less of a role once the norm has become widely known.

The second question that seems worthy of future investigation is why the rate of enforcement in our experiment is substantially lower relative to that in laboratory experiments. Presumably, there are many reasons. The absence of counter-punishment opportunities in most laboratory experiments is clearly an important reason as suggested by our survey evidence. However, this factor cannot account fully for the difference in enforcement rates as the majority of individuals seem to still use costly punishment when counter-punishment is possible in the lab (Nikiforakis, 2008; Nikiforakis and Engelmann, 2011). Based on our surveys, we believe that punishment may be considered to be more costly in the field than in laboratory experiments, as angry violators can cause physical or psychological damage to enforcers. It is also possible that individuals who violate norms are (or consider themselves to be) better at retaliating than those who do not. This also increases the cost of norm enforcement and may thus lower the frequency with which norm violations are punished. Another explanation for the higher rate of enforcement in the laboratory may be the fact that at the start of the experiment a clear social norm does not exist or has not been established and individuals may be more likely to use costly punishment when a norm is not well established. Understanding the reasons behind the difference in enforcement rates in the field and in the laboratory will yield insights into the determinants of costly punishment and how it can support norms of cooperation.

Finally, it will be interesting to examine individuals' willingness to engage in costly norm enforcement in other natural field experiments across different populations and across different social norms.

## References

- Anderson, C. & Putterman, L. (2006) Do Non-strategic Sanctions Obey the Law of Demand? The Demand for Punishment in the Voluntary Contribution Mechanism. *Games and Economic Behavior* 54, 1–24.
- Carpenter, J. (2007) The Demand for Punishment. *Journal of Economic Behavior and Organization* 62, 522–542.
- Benson, P.G., “The Hawthorne Effect,” in W.E. Craighead and C.B. Nemeroff (eds.), *The Corsini Encyclopedia of Psychology and Behavioral Science* (New York: Wiley, 2000; Third Edition, Volume 2).
- Cabral, L., Ozbay, E. & Schotter, A. (2011) Intrinsic and Instrumental Reciprocity: An Experimental Study. *Mimeo*.
- Chaudhuri, A. (2011) Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. Forthcoming in *Experimental Economics*.
- Casari, M. (2011) Weak Reciprocity Alone Cannot Explain Peer Punishment. Forthcoming in *Behavioral and Brain Sciences*.
- Coleman, J. (1990) *Foundations of Social Theory*. Cambridge MA: Harvard University Press.
- Denant-Boemont L., Masclet D. & Noussair C. (2007) Punishment, Counterpunishment and Sanction Enforcement in a Social Dilemma Experiment. *Economic Theory* 33, 145-167.
- Egas, M. & Riedl, A. (2008) The Economics of Altruistic Punishment and the Maintenance of Cooperation. *Proceedings of the Royal Society B* 275, 871–878.
- Elster, J. (1989) Social Norms and Economic Theory. *Journal of Economic Perspectives* 3, 99-118.
- Edwards, R. (2011) Police Officers Beaten by Mob after Asking Girl to Pick up Litter. *The Telegraph*, September 6. Available at: <http://www.telegraph.co.uk/news/uknews/2433514/Police-officers-beaten-by-mob-after-asking-girl-to-pick-up-litter.html>
- Fehr, E. & Fischbacher, U. (2002) Why Social Preferences Matter – The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives. *Economic Journal* 112, C1-C33.

- Fehr, E., & Fischbacher, U. (2004) Third-party Punishment and Social Norms. *Evolution and Human Behavior*, 25, 63-87.
- Fehr, E. & Gächter, S. (2000) Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90, 980-994.
- Fehr, E. & Gächter, S. (2002) Altruistic Punishment in Humans. *Nature* 415, 137-140.
- Fresco, A. (2008) Four Years for Teenagers who Killed Evren Anil. *The Times*, April 5.  
Available at: <http://www.timesonline.co.uk/tol/news/uk/crime/article3684893.ece>
- Fudenberg, D., Levin, D.K. & Maskin, E. (1994) The Folk Theorem with Imperfect Public Information. *Econometrica* 62, 997-1039.
- Fudenberg, D. & Maskin, E. (1986) The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica* 54, 533-54.
- Gächter, S. & Herrmann, B. (2009) Reciprocity, Culture, and Human Cooperation: Previous Insights and a New Cross-cultural Experiment. *Philosophical Transactions of the Royal Society B -- Biological Sciences* 364, 791-806.
- Guala, F. (2011) Reciprocity: Weak or Strong? What Punishment Experiments Do (and Do Not) Demonstrate. Forthcoming in *Behavioral and Brain Sciences*.
- Harrison, G.W. & List, J.A. (2004) Field Experiments. *Journal of Economic Literature* 42, 1009-1055
- Herrmann, B., Thöni, C. & Gächter, S. (2008) Antisocial Punishment Across Societies. *Science* 319, 1362-1367.
- Hoffman, E., McCabe K. & Smith, V. L. (1996) Social Distance and Other-Regarding Behavior in Dictator Games. *The American Economic Review* 86, 653-660.
- Hopfensitz, A. & Reuben, E. (2009) The Importance of Emotions for the Effectiveness of Social Punishment. *Economic Journal* 119, 1534-1559.
- Levitt, S. D. & List, J. A. (2007) What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World, *Journal of Economic Perspectives* 21, 153-174.
- Masclet, D., Noussair, C., Tucker, S. & Villeval M.C. (2003) Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism. *American Economic Review* 93, 366-380.

- Nikiforakis, N. (2008) Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves? *Journal of Public Economics* 92, 91–112.
- Nikiforakis, N. & Engelmann, D. (2011) Altruistic Punishment and the Threat of Feuds. forthcoming *Journal of Economic Behavior and Organization* 78, 319–332.
- Nikiforakis, N. & Normann, H.T. (2008) A Comparative Statics Analysis of Punishment in Public-Good Experiments. *Experimental Economics* 11, 358–369.
- Noussair, C., van Soest, D. & Stoop, J. (2011) Punishment, Reward, and Cooperation in a Framed Field Experiment. *MPRA Paper 34067*.
- Reuben, E. & Suetens S. (2011) Revisiting Strategic versus Non-Strategic Cooperation. Forthcoming in *Experimental Economics*.
- Spiegel, Der (2008) Immigrants Get Long Sentences For Attacking German Pensioner. *Der Spiegel*, August 7. Available at:  
<http://www.spiegel.de/international/germany/0,1518,564651,00.html>
- Spiegel, Der (2010) Murder on a Station Platform: German Teenagers Jailed for Killing 'Hero'. *Der Spiegel*, June 9. Available at:  
<http://www.spiegel.de/international/germany/0,1518,715963,00.html>
- Sugden, R. (1986) *The Economics of Rights, Cooperation and Welfare*. Oxford: Basil Blackwell.
- Xiao, E. & Houser, D. (2005). Emotion Expression in Human Punishment Behavior. *Proceedings of the National Academy of Sciences of the USA*, 102(20), 7398-7401.
- Young, H. P. (1998) Conventional contracts. *Review of Economic Studies*, 65, 773-792.
- Young, H. P. (2008) *Social Norms*. The New Palgrave Dictionary of Economics. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.



## Appendix

**Table A1 – Enforcement rates in Herrmann, Thöni and Gächter (2008)<sup>a</sup>**

City	% of subjects punishing at least once	% of punishment cases aimed at free riders <sup>b</sup>	% of subjects punishing at least once, periods 6-9	% of punishment cases aimed at free riders <sup>b</sup> , periods 6-9
Athens	88.6	62.8	79.5	63.3
Bonn	78.3	69.8	45.0	63.5
Boston	58.9	72.8	17.9	86.2
Chengdu	90.6	72.8	54.2	70.2
Copenhagen	67.7	64.3	27.9	58.5
Dnipropetrovs'k	86.4	63.0	77.3	67.4
Istanbul	89.1	62.6	67.2	65.2
Melbourne	82.5	69.9	50.0	63.1
Minsk	76.5	58.7	55.9	56.5
Muscat	90.4	49.3	76.9	50.2
Nottingham	82.1	76.5	44.6	66.7
Riyadh	79.2	59.4	56.3	56.9
Samara	89.5	61.2	75.7	64.7
Seoul	84.5	68.1	58.3	68.8
St. Gallen	79.2	66.0	44.8	65.3
Zurich	82.6	69.4	51.1	69.5

<sup>a</sup> Note that Herrmann et al. (2008) use a fixed matching protocol. However, while the matching protocol sometimes affects the extent of punishment (conditional on punishment being meted out), the *rate* of punishment, which is of interest to us, does not typically differ across fixed and random matching protocols. For some supporting evidence see Table 4 in Nikiforakis (2008) and the evidence discussed in section 2.

<sup>b</sup> The term “free rider” refers to individuals who contributed less than the average of the group.