The Effects of Increasing the Certainty of Punishment: A Field Experiment on Public Transportation
Martin Killias, David Scheidegger and Peter Nordenson
European Journal of Criminology 2009 6: 387
DOI: 10.1177/1477370809337881

The online version of this article can be found at:
http://euc.sagepub.com/content/6/5/387
The Effects of Increasing the Certainty of Punishment
A Field Experiment on Public Transportation

Martin Killias
University of Zurich, Switzerland

David Scheidegger
Zürcher Verkehrsverbund (ZVV; Zurich Transport Network), Switzerland

Peter Nordenson
Zürcher Verkehrsverbund (ZVV; Zurich Transport Network), Switzerland

ABSTRACT

Since 1993, Zurich’s suburban transport systems have operated without attendants on the trains, and passengers’ tickets were checked only sporadically. After increasing worries about crime and passenger safety on suburban trains, it was decided, in order to reduce fear of crime, to reintroduce attendants on all trains after 9 p.m. As well as dealing with order on the trains, these attendants had to check the tickets of all passengers. The programme was launched in June 2003, and once it was fully implemented the resources devoted to ticket checks and, thus, the number of passengers controlled after 9 p.m. rose by roughly 1500 percent. However, ticket checks remained stable and sporadic during day-time hours. The programme led to a dramatic reduction in fare-dodging on trains operating during the evening and, as an unexpected outcome, during day-time hours. The decline in fare evasion eventually levelled off. The results suggest that certainty of punishment works as a deterrent in a non-linear way, and that benefits from increased certainty can be maximized if checks are concentrated on critical hours and areas.

KEY WORDS

Certainty of Punishment / Deterrence / Fare Evasion / Natural Experiments / Public Transportation.
The effects of sanctions on human behaviour have been debated for centuries. The first formulation of the theory of deterrence dates back to Platon’s dialogue *Protagoras* (5th century BC) in which Socrates suggested that punishment should be imposed not so much for retribution, but mostly in order to prevent future crimes.\(^1\) Two generations later, Aristotle suggested, in his work on *Rhetorics*,\(^2\) that crime is more likely to be committed when the likelihood of punishment is low, or the benefits generated by the offence exceed the expected costs of the sanction. Modern writers from Beccaria to von Pufendorf, von Feuerbach, Bentham and Andenaes have suggested formulations of the theory that are strikingly similar, the two key variables remaining (since Aristotle) the severity and the certainty of punishment. Cesare Beccaria may have been the first scholar to have fully realized the importance of the probability (and not just of the severity) of the sanction.\(^3\) Today, the certainty of punishment is usually considered more important than severity in deterring offences (e.g. Doob and Webster 2003).

There are several historical examples where dramatic variations in the certainty of punishment have been followed by substantial increases (or decreases) in crime rates. For example, the dismantling of the Danish police in 1944 by the German occupation forces produced a sharp increase in several predatory crimes – until order was restored after the war (Trolle 1994; Andenaes 1974). More recent research, however, has generated rather contradictory results (e.g. Pratt et al. 2006). Campaigns aimed at curbing drunk-driving or other traffic offences have mostly been followed, at best, by moderate and temporary (rather than permanent) reductions in alcohol-related crashes (Ross 1982). Deterrent effects are often short-lived (Nagin 1998). The underlying problem may be that deterrence is not operating in a linear way, as Zimring and Hawkins (1973: 347) have recognized, and that the certainty (as well as the severity) of a sanction must reach levels where variations in the probability of imposition are likely to shape subjective risk assessments. If, for example, the risk of a driver being stopped for driving under the influence of alcohol is about 1 in 2000 (as estimated by Kerner 1985 for German drivers), any police campaign that doubles that risk may fail to increase subjective probabilities in any significant way (Yu and Liska 1993). Similarly, crackdowns against drug dealers or other types of offenders often affect subjective probabilities of an arrest far too little to produce any discernible deterrent effect.

Based on a systematic review of research, Beyleveld (1980: 308) concluded that substantial deterrent effects can be expected between a lower and an upper threshold (of certainty or severity) only. However, it will not

---

1. *Protagoras* (324, a, b).
3. In his landmark text *Dei delitti e delle pene* (Beccaria 1764: §§ 19 and 27).
be easy to observe empirically at what levels of objective probability (or severity) any additional increment of threat will, among potential offenders, increase their awareness of risk and thus reduce offending rates. In practice, the occasions will be rare where continuous increases in the certainty of punishment can be observed over time, since interventions will usually consist of one unique change in the independent variable. Therefore, most evaluations are limited to comparing crime rates before and after a unique intervention.

The natural experiment reported in this article offers an opportunity to observe the effects of continuous changes in the certainty of sanctions. The dependent variable was the number of passengers using the Zurich regional public transport system without a valid ticket and was measured monthly over several years. Over a period of more than one year, the intensity of passenger checks was steadily increased and reached, at the end of the period, roughly three times pre-intervention levels. During evening hours, passenger checks were multiplied by a factor of about 30. During the entire period, day-time checks of passengers remained unchanged.

The Zurich public transportation natural experiment

As was the case with other urban public transport systems in Europe, over the past 20 years the Zurich metropolitan transport company, which operates in an area that extends for more than 50 km from the centre of Zurich, has reduced the number of staff and ticket checks on regional trains in order to save costs. From 1993, all suburban trains operated without ticket checks. As a result, an increasing proportion of users started to evade payment of fares (Hauber 1993). According to internal data, approximately 4 percent of passengers were regularly found without a valid ticket. This rate remained stable over many years. Different measures taken to curb fare-dodging (such as checks by plain-clothes agents or by patrols of several agents acting together) failed to produce permanent reductions in fare-dodging.

Passengers found without a valid ticket receive an ‘on the spot’ fine. The fine was originally 60 Swiss francs (or approximately €40) in cash, or SFr 80 (or €54) in the case of deferred payment. From 1 December 2003, the amount for payment in cash was raised to SFr 80 (or €54). Under this system, passengers were not required to prove their identity. From the beginning of 2006, passengers in all cases were required to prove their identity and to pay SFr 80 in the case of a first offence. In the case of a third offence within two years, the amount increases to SFr 120 and criminal prosecution will be initiated. Passengers who are unable to prove their identity or who do not cooperate are arrested by the police. In these cases, as well as in the case of
repeat offenders, criminal prosecution will be the rule, usually leading to a conviction under section 150 of the Criminal Code\(^4\) and to a fine. As Figure 1 illustrates, the rate of fare evasion remained at nearly 4 percent despite increasing fines and changes in enforcement policies.

The natural experiment reported here was, however, aimed not at deterring fare evasion but at reducing the fear of crime among passengers. The removal of attendants has indeed created new opportunities for offending. Surveys conducted in 2000 and 2001 revealed that 54 percent of female passengers and 19 percent of male passengers felt unsafe while riding on evening trains (Killias and Lamon 2002). Fear among passengers was significantly more widespread than fear on the streets (after 10 p.m.), where 35 percent of women and 10 percent of men who regularly use public transport report feeling unsafe. Higher fear levels among passengers are not necessarily irrational, since respondents using public transport have a 60 percent (men) to 70 percent (females) higher risk (odds ratio) of personal victimization than non-users, once several other variables are controlled for (Killias 2002: 306). Regular surveys conducted by the Zurich Public Transport Network between 2000 and 2006 showed a substantial deterioration in passengers’ perceptions of safety after 9 p.m., particularly after a series of incidents reported in the media during 2002. Pressed by local parliaments, the metropolitan transport company subsequently decided to reintroduce attendants on all evening trains after 9 p.m. The purpose of this change was, therefore, to increase safety on evening trains and to reduce vandalism and disorder on the trains. However, the company took advantage of this change to reintroduce, at the same time, systematic ticket checks of all passengers during evening hours. Therefore, the reintroduction of attendants led to a substantial increase in the checking of tickets on evening trains in the Zurich region.

The reintroduction of attendants on all evening trains and the hiring and training of some 200 staff could not be implemented overnight, but took more than two years. Figure 1 illustrates how the new policy of increasing the number of hours devoted to checks was implemented on all 11 lines of S-trains operating in the Zurich region. As the programme progressed, the number of controlled metropolitan lines increased gradually over this period and the programme became fully implemented after 30 months. Since then, virtually all trains leaving from Zurich have been subject to ticket checks after 9 p.m. With two attendants on regional trains, composed of three double-level coaches containing 400 seats, it is assumed that roughly one passenger in three is checked during evening hours. During the same period,

\(^4\) This section criminalizes the evasion of fares or entry fees in all kinds of circumstances (public transport, theatres, etc.). In practice, most offenders are sentenced to pay a fine.
there was no change in fines or in the type or frequency of checks during the day, or on buses or trams operated by the Zurich public transport agency.

Results

In order to assess the effects of increased levels of checks on fare evasion on Zurich’s suburban public transport system, the number of working hours devoted to ticket checks was used as an independent variable. Working hours could be differentiated between day-time (spot checks) and evening hours (i.e. after 9 p.m.). It can be assumed that the number of passengers actually affected is highly correlated with variations in the working hours assigned to this task. The dependent variable is the percentage of passengers on trains found without a valid ticket. Before 2006, this percentage is known only for all hours of operation combined (Figure 1). From January 2006, this percentage is known for evening hours (with train attendant) as well as for day-time hours (with spot checks) separately, as well as for all hours of operation combined (Figure 2).

The outcomes will be presented separately for the period from early 2003 to the beginning of 2006 (Figure 1), and for the entire year 2006 (Figure 2). As shown in Figure 1, the number of working hours devoted to sporadic ticket checks during day-time hours remained constant between 90,000 and 100,000 during the entire period (left scale). However, checks during evening hours (i.e. on trains operating after 9 p.m.) increased from less than 10,000 to nearly 60,000 working hours during June 2003, and remained constant at this level until March–April 2004, when they increased to more than 130,000 (Figure 1, left scale). In January 2006 the number of working hours devoted to ticket checks on evening trains increased again, reaching a maximum of nearly 180,000 (Figure 1, left scale). In April–May 2006 the number of working hours devoted to ticket checks on evening trains decreased to less than 160,000 (Figure 2, left scale). How did this massive variation in checks on evening trains affect fare-dodging?

As one can see, the proportion of passengers travelling without a valid ticket on any train of the Zurich metropolitan transport system remained fairly stable (oscillating around 3.5 percent) between the beginning of 2003 and August of the same year. From September 2003, i.e. about two months after the number of ticket checks on evening trains had increased by about 600 percent, however, the rate of fare-dodging started to drop until May 2004, when it levelled off at about 1.2 percent. After March 2004, the number of working hours devoted to ticket checks on evening trains increased again, but fare-dodging stabilized finally at about 1.0 percent at the end of 2005.
Figure 1  Working hours devoted to ticket checks in the Zurich suburban public transport system, during day-time and evening hours (left scale), and passengers found during ticket checks without a valid ticket (PWVT, in percent of all passengers checked), all operating hours (right scale), from January 2003 to March 2006.
Figure 2  Working hours devoted to ticket checks in the Zurich suburban public transport system, during day-time and evening hours (left scale), and passengers found during ticket checks without a valid ticket (PWVT, in percent of all passengers checked), during daytime and evening hours as well as during all operating hours (right scale), from January to December 2006.
The additional increases in working hours devoted to ticket checks during the evening early in 2006 and the following reductions in April–May 2006 do not seem to have affected the rate of fare-dodging to any major extent (Figure 2). Fare-dodging is, as the more differentiated measures for day-time and evening hours given in Figure 2 show, slightly more frequent on day-time trains (with sporadic checks) than during evening hours (with systematic passenger checks). However, both lines are perfectly parallel, and there is no reason to suspect that the overall measure (for all hours of operation combined) masks important differences in trends between trains operating during the day and in the evening.

Overall, the rate of fare-dodging on Zurich regional trains has dropped by about 80 percent since the start of the programme. This was owing to the fact that checks increased very substantially, i.e. from about 100,000 hours early in 2003 to about 270,000 early in 2006 over the entire time of operation, and by a factor of between 6 and 20 during evening hours. This inflated deterrence was sufficient to reduce fare-dodging dramatically. The data do not allow us to specify when fare-dodging started to decrease, since the increase in checks occurred very suddenly.

The data in Figure 1 carry a second message. Beyond an increase of resources devoted to passenger ticket checks of about 100 percent overall, and a factor of six during evening hours, the marginal benefits (in the form of additional reductions in fare evasion) started to level off, and no decrease was observed once checks increased again in early 2006. Thus, there seems to be a ceiling beyond which additional increments in certainty seem no longer to pay off. Passengers who are (still) not discouraged by dramatically increasing levels of control may have different personality and social characteristics, and their decision to offend may be hard to overcome by any policy changes. They may have decided to offend despite the high risk, just as others may abstain from offending despite low probabilities of detection (Diekmann 1980; Künzli 1996). Another explanation could be that certain people with a risky lifestyle or on welfare are less able to adjust (or are less vulnerable) to increased risks – even if the increase is dramatic and easily visible.

**Diffusion of benefits**

As mentioned, the programme was designed to reduce the fear of crime on evening trains in the Zurich region, and not to discourage fare-dodging. Despite that, the programme had a dramatic effect on the prevalence of fare evasion. Interestingly, this reduction was observed during day-time as well as during evening hours, although the systematic rather than sporadic
checking of passengers was limited to evening trains operating after 9 p.m. The increase in resources devoted to ticket checks was also concentrated exclusively on evening trains, whereas day-time checks did not change in frequency or in style. In purely rational terms, one might have expected that passengers who are committed to fare-dodging would simply return home before 9 p.m., i.e. before systematic checks start.

The reason this did not occur may be found in the high prevalence of fare-dodging among young people. According to data from the Zurich regional transport system,\(^5\) 48 percent of passengers found without a valid ticket are under 26, although this age group represents only 27 percent of passengers. Since young people frequently use trains, buses and trams during the evening in connection with leisure-time activities, they probably do not wish to return home before 9 p.m. On the other hand, season tickets that offer free travel all day long are available for persons under the age of 26 at fairly low rates. Therefore, it would be far more rational (in economic terms) to buy a season ticket and have full access to the public transport system than to buy (relatively expensive) tickets for single trips during the evening and to continue travelling without a ticket during the day.

Discussion: What lessons can be learned from this experiment?

Any conclusions on the deterrent effects observed in Figure 1 are contingent on the validity of the indicators used. As far as numbers of working hours are concerned, there is no evidence regarding the validity of this measure as an indicator of the number of passengers controlled, although it is intuitively hard to believe that these two variables should not be highly correlated. The percentage of passengers found during checks travelling without a valid ticket may be of more questionable validity as an indicator of fare-dodging, just as offences known to the police never perfectly measure the actual numbers of crimes committed. In the present context, however, the natural experiment’s internal validity depends not on the precise measure of the ‘true’ volume of fare-dodging but on whether or not the proportion of passengers found without a valid ticket by attendants reflects trends in this variable over time. As a proportion, the validity of this indicator remains unaffected by changing numbers of passengers (which increased by 31 percent

\(^5\) The high prevalence of fare evasion among young people is well documented in Swiss self-report studies (Killias et al. 1994; Eisner et al. 2000; Ribeaud and Eisner 2009). One of the reasons is the high density of public transportation in Switzerland, and the fact that young people travel to school as well as to leisure-time activities by train, tram and bus far more often than young people in other countries.
between 2000 and 2005). Since the ‘styles’, tactics and modalities of checks, as well as the sanctions, have remained unchanged over the critical period, there is no reason to believe that changes at any of these levels (rather than a real change) might account for the dramatic drop. Nor is it plausible that the dramatic drop in fare-dodging was a result of higher security standards and lower fear levels among women and elderly people who, as a consequence, might have increasingly used public transportation after the programme started.

Further confirmation comes from two surveys on self-reported delinquency among 15-year-old students at schools in the Zurich area conducted in 1999 and in 2007. According to this source, the life-time prevalence of fare-dodging decreased from 62 percent to 52 percent between the two sweeps (Ribeaud and Eisner 2007). In sum, there is no reason to believe that the drop in fare-dodging seen in Figure 1 is an artefact.

As in many cases of ‘unplanned’ natural experiments, it is unfortunate that data on certain potentially crucial variables have not been collected. For example, it would have been helpful to have had data on the drop in fare evasion in 2003 separately for evening and day-time hours. It would also have been interesting to examine data on the subjective certainty of detection among passengers. Finally, it would have been helpful to follow trends separately for passengers in several age groups. Nevertheless, we think that the drop has been sufficiently substantial to justify some sound conclusions. For example, there is good reason to believe that the re-introduction of attendants on Zurich’s regional trains during the evening has led to a significant drop in the number of passengers travelling without a ticket. However, the decline in fare-dodging started to level off after the first (huge) increase in human resources devoted to checking passengers. Thus, the present experiment offers additional support to Beyveld’s (1980: 308) assertion that deterrent effects are not linear and that the marginal benefits of increasing checks tend to reach a ceiling. If this conclusion is accepted, checks (and, thus, the certainty of punishment) should not be maximized, but an optimal level should be sought if some balance between investments in control and returns in the form of additional ticket sales is of concern.

The results of this natural experiment show, in line with many others (Farrington and Welsh 2006), that diffusion of benefits is far more likely to occur once offending is made more difficult in certain places or, as here, during certain hours. If systematic ticket checks during certain evening hours are reducing fare evasion on day-time trains as well, it might be an economic and efficient option to concentrate checks on certain critical hours, and to push persistent fare dodgers into buying season tickets as the most economic alternative. At the extreme, one might imagine (and eventually test)
systematic passenger checks only on last trains or on trains operating during the last hour of the day. An alternative option might be to target checks on morning trains from, say, 7 to 9 a.m., considering that most young people are obliged to travel to school during these hours. On the other hand, and owing to the enormous passenger traffic during these hours, checks in the morning are probably far more expensive to operate than checks during the less busy evening hours. Of course, this is an option only in countries where many young people use public transportation for evening activities.

In addition, checks during evening hours may also produce benefits beyond reducing fare evasion. For example, harassment of passengers, violence and vandalism are all concentrated during evening hours and may be reduced through ticket checks on late evening trains. In the present example, combating fear of crime was the first priority when the programme started, and internal survey data6 show that the percentage of passengers feeling at least fairly safe while travelling on suburban trains increased between 2002 and 2006. In addition, data on damage to suburban trains show a reduction by 18 percent between 2004 and 2006.

Thresholds also seem to play a role in other domains where, beyond punishment, people tend to avoid negative consequences. Fear of crime, for example, does not consistently increase or decrease according to changing crime trends. However, if crime rates vary dramatically, for example across neighbourhoods, fear levels tend to correlate strongly with objective risks.7 The same is true of campaigns calling on the public to take preventive action against certain risks such as diseases: again, such campaigns may affect behaviour only if the risks are perceived to increase very substantially,8 and, as in the case of fare-dodging, there seems to be a fraction of individuals not amenable to more responsible behaviour. In sum, we may face great difficulties in identifying changing risk levels as long as variations are not dramatic.

6 D. Scheidegger, Sicherheitskennzahlen Zürcher Verkehrsverbund 2006.
7 Studies in Switzerland (Killias 2002: 406–7) as well as in the Netherlands (Wittebrood 2000) have shown very high correlations between rates of personal victimization and fear of crime across neighbourhoods. In both countries, the risks of personal crime are up to 10 times higher in ‘bad’ neighbourhoods compared with affluent residential areas. Variations across space or time are usually far less extreme, and varying risk levels may, therefore, less easily translate into corresponding levels of fear.
8 Evaluations of campaigns designed to prevent AIDS (and to promote the use of condoms) have shown very high rates of behavioural change, including a decline in the proportion of respondents not using condoms from 67 to 13 percent within 20 months (IPSO 1991, in Killias 2002: 445).
Conclusions

The Zurich natural experiment on fare-dodging revealed a substantial and lasting reduction in fare-dodging on suburban trains after ticket checks were reintroduced on virtually all trains operating after 9 p.m. This success was not limited to evening trains but extended to all hours of operation (owing to the relatively high costs of single fares and the rather low prices of season tickets). It also turned out that, beyond a certain upper threshold, additional increments of certainty of detection did not reduce fare evasion beyond a certain level, suggesting that some people may remain unaffected by increasing threats. This finding adds to an increasing body of research suggesting that, whenever people tend to avoid negative consequences, they react only to extreme variations in risks.

Although the external validity of any natural or controlled experiment remains to be seen, the observation of a ceiling effect, i.e. an upper threshold beyond which increasing certainty levels no longer affect behaviour, is important. Since occasions are rare where the critical independent and dependent variables can be measured over longer periods of time, the results presented here may be relevant far beyond the prevention of fare-dodging. They also demonstrate that displacement effects are far less likely than diffusion of benefits.

References

The effects of increasing the certainty of punishment


---

**Martin Killias**

Martin Killias is Professor of Criminology and Criminal Law at the University of Zurich Law Faculty. He has initiated several randomized and natural experiments in the field of deterrence and corrections. His main work focuses on national and international surveys concerning victimization and self-reported delinquency. He also chairs the European Sourcebook of Crime and Criminal Justice Statistics project.  
martin.killias@rwi.uzh.ch
David Scheidegger

David Scheidegger works at the Zürcher Verkehrsverbund (ZVV) as a quality representative. He studied at the Swiss Federal Institute of Technology Zurich and gained a doctorate in chemistry. From 1997 he worked as an independent consultant on the environment, quality management and safety at work. In 2001 he was appointed training and quality manager of a company in the facility services sector. At the end of 2002 he moved to ZVV to focus on service quality, security and quality agreements.

David.Scheidegger@ZVV.zh.ch

Peter Nordenson

Peter Nordenson is responsible for Marketing Control at the Zürcher Verkehrsverbund (ZVV). He has an MSc in Transportation and Logistics from Chalmers University of Technology in Sweden and worked for several years developing logistics and marketing strategies for an international consumer goods company, before joining ZVV in 1997.