Discretion and Ulterior Motives in Traffic Stops: The Detection of Other Crimes and the Revenue from Tickets*

Brian Rowe†
Bureau of Economics
Federal Trade Commission
browe@ftc.gov

April 2010

Abstract

In the United States, police officers often decide to give drivers they stop for traffic violations a warning, which imposes no fine, instead of a ticket. Officers are also legally permitted to stop drivers for the purpose of detecting other crimes. This paper addresses two questions about the role of discretion and ulterior motives in traffic stops. First, under what conditions may it be efficient to let many stopped drivers go with only a warning? Using a model of law enforcement based on Shavell (1991), who does not consider warnings, I show that the ulterior motive of detecting other crimes is a simple way to rationalize the existence of warnings in an efficient enforcement scheme. Second, I test the model against data on traffic tickets and warnings in Massachusetts to determine whether police discriminate against out-of-town drivers because of the ulterior motive of ticket revenue. I find support for the notion that discrimination against out-of-town drivers is motivated by revenue. In the model, the revenue motive is an aspect of efficient enforcement for a local government.

*I am grateful to Paul Courant, Jim Hines, Stephan Lindner, Yusufcan Masatlioglu, J.J. Prescott, Paul Rothstein, Matt Rutledge, Dave Schmidt, Dan Silverman, Doug Smith, Jeff Smith, and seminar participants at the University of Michigan, and the Association for Public Economic Theory’s 2009 meeting for many helpful comments. I thank Kate Antonovics, Bill Dedman, and Nicola Persico for sharing data.

†The views expressed in this article are those of the author and do not necessarily reflect those of the Federal Trade Commission.
1 Introduction

Police officers in the United States practice a “catch and release” policy, because they decide to let many of the drivers they stop for traffic violations go without being issued a ticket. According to recent Bureau of Justice Statistics data, 40.3% of stopped drivers in the U.S. were not required to pay a fine¹, an outcome which I refer to as a warning. This use of discretion in whether to give stopped drivers a warning raises several important issues. First, does the high prevalence of warnings imply that traffic enforcement is grossly inefficient? Sizable public resources are expended to detect, stop, and prosecute violators of traffic laws.² Traffic laws are intended to promote public safety on the roads, and traffic accidents are a significant public health problem. Approximately 42,000 people were killed and 2.5 million people were injured in traffic accidents in 2006 (NHTSA 2007). Yet because many stopped drivers are only given warnings, the expected cost to drivers of committing a traffic violation is below what it would be if all stopped drivers were ticketed. As a result, less deterrence of traffic violations, and therefore more traffic accidents, may be realized for a given level of public expenditures on traffic enforcement.

Second, the impact of police discretion in traffic stops has generated a great deal of controversy with respect to racial profiling and racial disparities in searches of stopped motor vehicles for drugs.³ The ability to give tickets or warnings provides another way in which the police can take different actions against drivers who committed identical traffic violations. Discretion in ticketing may therefore lead to unjustified differential treatment in ticketing based on characteristics such as race and gender.⁴ In light of these concerns, an important and policy relevant question is whether there is any good reason to allow the police to issue warnings. This paper provides an answer, based on the importance of the ulterior motive of detecting other crimes.

In the U.S., the police are legally permitted to stop drivers for minor traffic violations for the purpose of detecting other criminal activities. In Whren et al. vs United States 1996, the Supreme Court held that is legal for police officers to stop drivers for minor traffic violations if they suspect the driver is guilty of another crime. Police may do this even if a “reasonable” officer would not have stopped the driver for the traffic violation alone.⁵

¹These drivers received either a written warning, a verbal warning, or nothing at all (Durose et al. 2007).
²About 17.8 million drivers were stopped by the police in 2005 (Durose et al. 2007), and court cases related to traffic violations accounted for more than half of all cases in state courts in 2005 (National Center for State Courts, 2007).
³See Harris (1997) for an influential argument that police discretion has an adverse impact on minority motorists. Ramirez et al. (2000) summarize many studies of racial profiling and document a dramatic increase in the collection of traffic stop data in response to concerns about profiling. According to Antonovics and Knight (2009), there have been over 200 court cases related to racial profiling in searches of stopped motorists.
⁴Dedman and Latour (2003) documented racial, gender, and age disparities in the probability of being ticketed after a stop in Massachusetts. In response, Massachusetts funded a follow-up study (Farrell et al. 2004), which reached similar conclusions.
⁵Holding of Whren: “The temporary detention of a motorist upon probable cause to believe that he has violated the traffic laws does not violate the Fourth Amendment’s prohibition against unreasonable seizures, even if a reasonable officer would not have stopped the motorist absent some additional law enforcement objective.”
The detection of other criminal activities is a key aspect of traffic enforcement. In 2005 about 4.8% of all stopped drivers were searched for drugs or other contraband, and 2.4% of all stopped drivers were arrested.\(^6\) Almost all of these searches and arrests occurred after the driver was stopped for a minor traffic infraction such as speeding or a vehicle defect. Since only 11.6% of searches uncovered evidence of criminal activity, many drivers were arrested for reasons unrelated to a vehicle search. Also, evidence from legal cases supports the idea that police commonly use minor traffic violations as a pretext for stopping motorists in order to look for other crimes.\(^7\)

Therefore, in order to analyze the conditions under which warnings might be efficient I employ a theoretical framework in which traffic stops accomplish a parallel objective of detecting other crimes. This framework is based on Shavell (1991) and Mookerjee and Png (1992), who formulated the idea of “general enforcement”, meaning that one detection probability applies to many different crimes.\(^8\) Using a general enforcement model based on Shavell (1991), who does not consider warnings, I show that when police stop drivers for the ulterior motive of detecting other crimes it may be optimal to let many stopped drivers go with only a warning. This holds despite the fact that in the model, the behavior of potential criminals is not affected by receiving a warning. Several factors combine to produce the result. First, when the enforcement authority can detect other criminal activity with traffic stops, more drivers are stopped for traffic violations than would be optimal if traffic offenses were the only crimes being monitored. Given fixed fines, if all the stopped drivers were ticketed the expected fine for minor traffic offenses could be higher than the social cost. The best way to reduce the expected fine to the optimal level (the social cost) is then to let a fraction of the stopped drivers go with only a warning. This preserves the benefit to society of potentially detecting more serious crimes, without over-detererring individuals from committing minor traffic violations.

This result provides a parsimonious justification for allowing police to give warnings which is consistent with empirical evidence about how and why traffic laws are actually enforced. Alternatively, warnings could be rationalized by assuming they have a direct deterrence effect, perhaps by reminding drivers to be more careful. I consider some implications of this explanation for warnings, but I find little support for them.

In practice, the police could abuse their discretion to give warnings. Instead of testing for discrimination by race or gender,\(^9\) I describe and analyze police discrimination against out-of-town

---

\(^6\)I arrive at 4.8% based on Durose et al. (2007), who estimate that 854,990 drivers were searched out of a total of 17.8 million stopped drivers. The percent of stopped drivers arrested is from Table 7 of the same report.

\(^7\)Summarizing racial profiling cases, Ramirez et al. (2000) state that “By far the most common complaint by members of communities of color is that they are being stopped for petty traffic violations such as under-inflated tires, failure to signal properly before switching lanes, vehicle equipment failures...” and so on. In this paper, I do not evaluate the costs and benefits of police discretion in choosing drivers to investigate for other crimes.

\(^8\)The main result of these two papers is that when the probabilities of detection for different crimes are linked, optimal fines rise with the severity of the crime, and maximal fines are only used for the most serious offenses.

\(^9\)Three papers do this using data from Massachusetts. Anbarci and Lee (2008) find that minority officers were less likely to reduce the charged speed for minority drivers. Antonovics and Knight (2009) find that mismatch between driver and officer race increases the probability of a search. Rowe (2010) develops a test for gender bias and finds that gender discrimination in ticketing exists but has only a small impact.
drivers in Massachusetts. This discrimination could result because towns in Massachusetts have “tax-exporting” incentives to collect additional revenue from out-of-town drivers. Articles about traffic enforcement in the popular press often describe a sense that the authorities may be more concerned with generating revenue than providing safe road conditions. Public opinion about revenue raising has even directly influenced enforcement policy. A backlash about the new revenue to be generated from fines forced the Virginia state legislature to retract a plan to substantially raise fines for reckless driving offenses (Craig 2008).

My analysis shows how a concern for revenue can be an aspect of an efficient enforcement policy for a local government. This depends on treating fine revenue from out-of-town drivers as a net transfer of wealth, and fine revenue from in-town drivers as a welfare-neutral transfer payment. After incorporating those conditions into the model, in the optimal policy warnings are given less frequently to stopped out-of-town drivers. In this way, a “revenue motive” behind ticketing is not necessarily in conflict with effective law enforcement.

Empirically, I find that local police in Massachusetts are more likely to ticket out-of-town drivers stopped for speeding than similar in-town drivers, controlling for a rich set of observable characteristics. This result alone, however, does not speak to why the discrimination occurs, because the police might treat out-of-town drivers more harshly for reasons unrelated to ticket revenue. I use the model to derive a prediction for an empirical pattern which should appear if police are motivated by ticket revenue. If motivated by revenue, the model predicts that discrimination against out-of-town drivers should disappear for more serious violations. I find support for this pattern in the data. At least with respect to the different treatment of out-of-town drivers, this suggests discretion in ticketing appears to be used efficiently.

1.1 Related Literature

Harrington (1988) considers a dynamic model based on the context of the enforcement of environmental regulations. The enforcement agency monitors firms repeatedly, fines and inspection probabilities can be conditioned on past compliance, and a maximum fine amount is exogenously fixed. Harrington (1988) finds that the optimal fine for “complying” firms is zero, as this provides the highest incentive for “violating” firms to comply with the law and thereby move into the “complying” group. My rationalization for warnings (a fine of zero for some violators) occurs in a static setting with no repeated interactions and no mechanism to condition enforcement on past compliance. Dittman (2006) analyzes a model where the government maximizes a weighted sum of social welfare and the “residual budget”, which is the revenue from fines net of enforcement expenditures. He shows that as the weight on the residual budget increases, the optimal probability of detection falls. Intuitively, the government induces more crime in order to catch and fine more criminals. Dittman excludes criminal utility from the welfare function because he focuses on serious crimes.

—

In my analysis, criminal utility is included and the concern for fine revenue arises only from the maximization of social welfare.

Three recent empirical papers address the question of how fine revenue influences enforcement. Baicker and Jacobson (2007) find that police agencies increase the drug arrest rate in response to the net fiscal incentives of asset forfeiture laws, which allow agencies to keep assets seized in drug arrests. Garrett and Wagner (2008) find that in North Carolina, lagged negative county level revenue shocks lead to small increases in the number of traffic tickets. Makowsky and Stratmann (2008) find that Massachusetts officers are more likely to issue a traffic ticket, and charge a higher fine amount, when their town’s budget is tight. They use probit models for whether stopped drivers were ticketed, and Heckman selection models of the fine amount charged to the driver.\textsuperscript{11}

The normative consequences of these revenue effects are ambiguous. Regarding Baicker and Jacobson, we do not know if the police increased their total enforcement effort or shifted their efforts towards drugs and away from other crimes in response to the seizure incentives.\textsuperscript{12} Regarding Garrett and Wagner, in times of fiscal distress it may be optimal for the county government to issue more traffic tickets if the alternative is to increase distortionary tax rates. The same argument applies to Makowsky and Stratmann’s result that towns with tight budgets set a higher probability of giving a ticket. My use of the general enforcement framework provides a clear prediction of patterns we should expect in the data if the government is concerned with fine revenue as part of an optimal policy.

2 Warnings and the detection of other crimes

2.1 Traffic stops only detect traffic violations

This section demonstrates that warnings are inefficient in a model where traffic laws are enforced only to deter traffic violations. The model is based on Polinsky and Shavell (1984), and I incorporate warnings. The following notation is used.

\( b \): Benefit from crime (breaking a traffic law).
\( G(b) \): Cumulative distribution function for \( b \).
\( x \): Public resources expended on detecting crime (monitoring the roads).
\( p(x) \): Probability of a violator being detected. \( p'(x) > 0 \).
\( F \): Fine amount, which is exogenously fixed.
\( c \): External cost of committing crime.
\( \pi \): Proportion of detected criminals (stopped drivers) who are fined.

\textsuperscript{11}To identify the Heckman model, an indicator for whether the driver had a commercial driver’s license is included in the equation for whether a ticket was issued but excluded from the fine amount equation.

\textsuperscript{12}Baicker and Jacobson provide evidence that police agencies shift their enforcement efforts towards heroin and away from marijuana in response to the budgetary incentives of seizure laws.
\[ z = \pi \times p(x) \times F \]: Expected fine from breaking the law.

Drivers make their decision about whether to break the law by comparing their benefit to the expected cost, and they are risk neutral so they violate the law if \( b > z \). Normalize the number of drivers to 1 so that the number of violators is equal to \( 1 - G(z) \), and then \( c[1 - G(z)] \) is the total social cost incurred from traffic infractions. The government chooses enforcement spending \( x \) to maximize the following social welfare function:

\[
\int_{p(x)F}^{\infty} bdG(b) - c[1 - G(p(x)F)] - x
\]

(1)

Equation (1) says the government takes into account the private gains drivers receive from violations (time saved by speeding), the social costs violators impose on others (increased risk of injury), and the cost of monitoring the roads. At the optimum \( x^* \), the following equation holds:\(^{13}\)

\[ p(x^*)F = c - \frac{1}{\frac{\partial G(z)}{\partial x}} \]

(2)

In the U.S. about 40% of stopped drivers are let go with only a warning. I now show that in the model just laid out, giving any warnings is inefficient. Consider \( \tilde{z} = \tilde{\pi} \times p(\tilde{x}) \times F \), which results in \( 1 - G(\tilde{z}) \) violators. If \( \tilde{\pi} < 1 \), then the enforcement authority could choose a \( \tilde{\pi} > \tilde{\pi} \) and an \( \tilde{x} < \tilde{x} \) such that the new expected fine \( \tilde{z} = \tilde{\pi} \times p(\tilde{x}) \times F \) is equal to the original expected fine \( \tilde{z} \). Since the expected fines are equal, the same number of drivers break the law and the social costs and benefits of crime are the same, but fewer resources are expended because \( \tilde{x} < \tilde{x} \). Therefore enforcement is inefficient if any drivers receive warnings. This is simply because the enforcement authority can costlessly substitute fewer warnings for reduced enforcement expenditures while maintaining the same level of deterrence.

The same conclusion will hold in a setting in which marginal deterrence is important. For example, suppose that drivers decide by how much to exceed the speed limit, and that faster violations are more dangerous. Given exogenous fines, the best use of resources by the enforcement authority would be to set a lower probability of stop for slower violations. Setting a higher probability of a warning for slower violations is a waste of resources because it is costly to stop the slow speeders in the first place.

2.2 Traffic stops detect other crimes

This section analyzes the case where the police use traffic stops to detect other crimes, using Shavell’s (1991) “general enforcement” model as the theoretical framework. I incorporate warnings into the model, whereas Shavell (1991) does not. In the model there are many different crimes,

\(^{13}\)The first order condition with respect to \( x \) is \( \frac{\partial p}{\partial x} F g(z) c - \frac{\partial g}{\partial x} F^2 g(z) p(x) = 1 \).
characterized by the different amounts of harm their commission imposes on society.
The elements of the model are as follows.

\( b \): Benefit from crime.
\( G(b) \): Cumulative distribution function for \( b \).
\( c \): External cost of crime.
\( S(c) \): Cumulative distribution function for \( c \).
\( x(c) \): Enforcement spending specific to detecting crime \( c \).
\( y \): General enforcement spending.
\( p(x(c), y) \): Probability of detection, increasing in both arguments.
\( F(c) \): Schedule of fines by crime \( c \), which is exogenously fixed.
\( \pi(c) \): Proportion of detected crime \( c \) criminals who are fined.
\( z(c) = \pi(c) \times p(x(c), y) \times F(c) \): Expected fine for committing crime \( c \).

Each individual in the model is characterized by the pair \( (b_i, c_i) \). Assume the distribution of benefits \( G(b) \) is the same for each crime \( c \). This guarantees that crimes with higher social costs are relatively more harmful to society.

In Shavell’s (1991) version of the model, he considers individuals as potentially able to commit only one crime. In contrast, I use the model to analyze law enforcement when people can potentially break traffic laws and commit more serious crimes. I do this in a simple way by assuming that such an individual appears in the social welfare calculation twice. This person is described by \( (b_1^i, c_1^i) \) and \( (b_2^i, c_2^i) \), where crime 1 might be speeding and crime 2 might be drug trafficking. Critically, this assumption implies that the individual’s choices of whether to commit traffic crimes and serious crimes occur independently of each other, a restriction which I justify shortly.

Unlike Shavell (1991), where fine amounts are choice variables, I assume that \( F(c) \) is fixed exogenously. This makes sense in the context of law enforcement by a local government because the schedule of fines is fixed by the state.\(^{14}\) Also, I assume that all fines \( F(c) \) must be less than some maximal fine \( \bar{F} \). The probability of detection for a given crime \( c \) is determined by the enforcement authority’s choices of spending specific to that crime, \( x(c) \), and how much to spend on general enforcement \( y \). For example, drunk driving is detected by police monitoring all activities on the roads, and by police checkpoints set up specifically to detect drinking drivers.

The government can also choose the probability of fining a detected crime \( c \) criminal, \( \pi(c) \). Therefore, the government’s problem is to choose \( \pi(c), x(c), \) and \( y \) to maximize the social welfare function \( W \):

\(^{14}\)Massachusetts state law Chapter 90, section 20 sets a schedule of fines for violating posted speed limits. Available at http://www.mass.gov/legis/laws/mgl/90-20.htm
\[ W = \int_{0}^{\infty} \int_{\pi(c) \times p(x(c), y) \times F(c)}^{\infty} (b - c) dG(b) dS(c) - \int_{0}^{\infty} x(c) dc - y \]  

The double integral in (3) is the net social benefit of crime. The two terms subtracted are the costs of enforcement. I assume that the optimal \( y \) (call it \( y^* \)) is positive, so that all crimes are detected with at least probability \( p(0, y^*) \).\(^{15}\)

The crucial restriction in this setup is that drivers who are also more serious criminals cannot alter their chances of detection by choosing how to violate traffic laws. This might seem unrealistic at first. However, the available evidence indicates that this restriction is actually appropriate in the context of traffic enforcement. Because traffic laws are numerous, complex, and often highly technical, it is almost impossible to avoid committing a violation that is grounds for a stop.

According to Harris (1997), a rule of thumb held by the police is that drivers cannot go more than three blocks without violating a traffic law. Ramirez et al. (2000) refer to a similar rule of thumb that drivers cannot go more than 1 or 2 minutes before a violation is observed. Harris (1997) also gives examples of laws that are sufficiently vague (such as that a driver may not “suddenly” slow down before turning) so that the police may invoke them as grounds for a stop at will. Knowles, Persico, and Todd (2001) similarly argue that because of Whren and the complexity of traffic codes, the legal constraint that police must have probable cause for a traffic stop is not binding in practice. Criminals may also find it difficult to reduce their chance of being detected by refraining from breaking traffic laws because most other drivers violate traffic laws. In a study which conducted direct observations of drivers on the New Jersey Turnpike, Lamberth (1994) found that 98% of drivers were exceeding the posted speed limit of 55 miles-per-hour. It seems unlikely that an intelligent drug trafficker would choose to drive exactly 55 in this situation. The police would notice that he was driving differently than everyone else, and then use some other traffic regulation as the grounds for a stop.

I now demonstrate that in this setting it may be optimal for the government to release some detected violators without requiring the payment of a fine. Letting \( * \) denote optimal values, I will show that \( 0 < \pi^*(c) < 1 \) for some crimes. The first step in the argument is:

**Proposition 1.** Warnings and specific enforcement cannot co-exist in an optimal policy for crime \( c \): If \( \pi^*(c) < 1 \), then \( x^*(c) = 0 \).

See the Appendix for the formal proof, but the intuition is straightforward. If any warnings were given for crime \( c \) and specific enforcement \( x(c) \) was positive, the same deterrence could be attained with fewer resources expended by increasing \( \pi(c) \) and decreasing \( x(c) \).

**Definition of minor crimes.** If \( x^*(c) = 0 \), then crime \( c \) is a minor crime.

\(^{15}\)Shavell (1991) shows that the optimal \( y \) in his formulation of the model will be positive if the average external cost \( E[c] \) is sufficiently large or if the maximal fine \( F \) is large.
Since warnings exist for most traffic crimes, we can say they are minor crimes. The detection probability for all minor crimes is \( p(0, y^*) \). Having chosen \( p(0, y^*) \), for minor crimes the local government can also choose \( \pi(c) \).

**Proposition 2.** For minor crimes, the optimal expected fine is equal to the external harm imposed: 
\[
z^*(c) = c, \text{ where } z^*(c) = \pi^*(c) \times p(0, y^*) \times F(c).
\]
Therefore if \( p(0, y^*) \times F(c) > c \), then \( 0 < \pi^*(c) < 1 \).

See the Appendix for the proof. As a first step to interpret the result, notice that the optimal general detection probability for minor crimes, \( p(0, y^*) \), takes into account the degrees of harm imposed on society by many different crimes simultaneously. When fine amounts are fixed by a higher level of government, if all detected criminals were punished it is possible that for some minor crimes the expected cost of breaking the law may exceed the harm imposed. If this occurs, the optimal way to reduce the expected fine for those minor crimes is to reduce the probability of fining people caught committing those crimes. This maintains the benefit of potentially detecting more serious crimes without over-deterring individuals from committing the minor crimes.

### 2.3 Discussion

This section explains several implications and limitations of the general enforcement framework for traffic stops. The first interesting implication is that when traffic stops are carried out with the ulterior motive of detecting other crimes, the enforcement authority stops more drivers for traffic violations than it would if traffic infractions were the only crimes being detected. To see this, compare the expected fine for a minor traffic crime \( c_1 \) under general enforcement with the expected fine when monitoring the roads only detects \( c_1 \).

\[
\begin{align*}
\text{Other crimes detected} & \quad : \quad \pi^*(c_1) \times p(0, y^*) \times F(c_1) = c_1 \\
\text{Only crime } c_1 \text{ is detected} & \quad : \quad p(x^*) \times F = c_1 - \frac{1}{\partial G[p(x^*)F]/\partial x} \\
\text{Assume the fines are equal} & \quad : \quad \implies \pi^*(c_1) \times p(0, y^*) > p(x^*)
\end{align*}
\]

For the same crime and equal fine amounts, the expected fine is higher under general enforcement. Therefore the probability of detection must also be higher. This makes sense in that all else equal, the ulterior motive of catching more serious criminals raises the benefit to the government of traffic stops. With fixed fines, some of the stopped drivers are released without a fine to avoid over-deterrence.

In the model, the optimal probability of receiving a ticket \( \pi^*(c) \) depends only on the general detection probability and the degree of harm imposed. This implies that within groups of stopped drivers who committed identical offenses, some will receive warnings and others will not. There

\[\text{16} \text{This is the solution given in equation (2).}\]
need not be other factors, observed by police but not by the researcher, to explain why some drivers are warned and others ticketed. In the context of the model, the police exercise discretion by randomly choosing some drivers for warnings. This contrasts with previous research on the use of discretion by government agents (such as prosecutors or police), where information that is private to those agents plays a central role. For instance, in Reinganum’s (1988) model prosecutor’s information about the strength of the case is the key variable in their decision of whether to dismiss the case, offer a plea-bargain, or take the case to trial. Shavell (2007) models the use of discretion as government agents basing their decisions on variables which they observe but which cannot be included in the law. A tradeoff results because given discretion, agents can base their decisions on factors which society views as relevant (such as extenuating circumstances or remorse), and also on factors which the agents care about but society has deemed irrelevant (such as race). Reinganum (1988) points out a different tradeoff by showing that restricting prosecutors to offer defendants charged with the same crime the same plea-bargain sentence is more likely to improve social welfare when the proportion of guilty defendants is higher.

Certainly there are potential costs of police discretion in ticketing that are not included in my model. For example, the implication that warnings exist within identical groups of drivers violates horizontal equity. People may have a preference for horizontal equity and therefore view this as undesirable. Using Shavell’s (2007) framework, another cost of discretion in traffic ticketing is that some decisions may be made on the basis of characteristics such as race or gender which are not relevant under the law. Although I do not explicitly model these costs, my analysis suggests that discretion in ticketing can be desirable simply if the benefit from detecting other crimes outweighs the various costs of police discretion. However, the model does not rule out a more active role for police discretion which takes into account information about the violation observed by the officer. Consider the traffic violation “Failure to Stop”, which is punished in Massachusetts with a $50 fine. Depending on the specific context, failing to stop can be relatively harmless or quite dangerous. If these offenses with different levels of harm are all minor crimes, they are all detected with the same probability $p(0, y^*)$. In this case, allowing the police to issue warnings more frequently for offenses which they think are less dangerous will avoid over-deterrence for less serious violations. Therefore, police discretion in giving warnings can be useful if the enforcement authority can choose a schedule of fines, but cannot perfectly condition the fines on the severity of the offense.

I have demonstrated that even if people only respond to the expected cost of breaking the law, when traffic stops detect other crimes it may be optimal to give warnings. A plausible alternative rationalization for warnings is simply that warnings have some deterrence value. For instance, receiving a warning might make a driver less likely to commit traffic violations in the future. If warnings have a deterrent value, we might expect warnings to be widespread in contexts where other crimes are not detected. Consider the case of automated traffic enforcement cameras, which take photographs of the license plates of vehicles which run red lights or exceed the speed limit.
Automated cameras make for a clean comparison to traffic stops; both detect traffic violations, but only traffic stops can detect other crimes. Speed cameras are widely used in the U.K., and while drivers may contest the fines in court, no warnings are issued.\(^{17}\) A speed camera program in Belgium tickets all drivers who exceed a certain threshold (Eeckhout, Persico, and Todd 2005).

According to the Insurance Institute for Highway Safety, red light cameras are authorized by state law and in use in 16 U.S. states, and speed cameras are used in 9 states.\(^{18}\) Colorado’s speed camera program is the only one which issues any warnings. In Colorado, first-time offenders caught exceeding the speed limit by less than 10 miles-per-hour are warned. This appears to be the sole exception. Clearly, warnings are much less prevalent when red light running and speed limits are enforced by cameras. This suggests that enforcement authorities see warnings as having little or no deterrence value.

Another relevant implication of the model is that if both cameras and traffic stops are used to detect the same crime (for example, speeding on a particular stretch of road), enforcement is inefficient if any of the traffic stops result in warnings. The cameras are specific enforcement because they only detect one crime, and by Proposition (1) it is inefficient to give warnings for a crime that is detected by both general and specific enforcement.

3 Data on tickets and warnings

In Massachusetts, police officers have legal discretion over whether to ticket a stopped driver or let him go with a warning.\(^{19}\) When a warning is given, the officer fills out the same information about the stop on a citation form, but checks a box at the bottom of the form. The data contain the information from the citation form for 166,368 traffic stops which occurred in Massachusetts in April and May of 2001. The sample is each traffic stop where a citation form was completed.

Some key variables in the data are the date and time of the traffic stop, the town where the stop was made, the driver’s hometown, and the reason for the stop. For speeding stops, both the speed limit and the miles-per-hour over the limit the driver was charged are available. The driver’s age, race, and gender are also included. Finally, the police agency which made the stop (Boston, state, or local police) is recorded.\(^{20}\)

\(^{17}\)See the Department for Transport’s website for a summary of the use of cameras at [http://www.dft.gov.uk/pgr/roadsafety](http://www.dft.gov.uk/pgr/roadsafety). An example of an effectiveness study is Christie et al. (2003).

\(^{18}\)The states which use speed cameras are Arizona, California, Colorado, Illinois, Maryland, Massachusetts, Ohio, Oregon, Tennessee, plus Washington D.C.. States which authorize and use red light cameras are the states above except for Massachusetts and Ohio, plus Delaware, Georgia, New Jersey, New York, North Carolina, Pennsylvania, Rhode Island, Texas, and Washington. The IIHS explains that cameras are sometimes used in municipalities in states where cameras have not been regulated by state statute. This information is from [www.iihs.org/laws/automated_enforcement.aspx](http://www.iihs.org/laws/automated_enforcement.aspx) and [www.iihs.org/research/topics/auto_enforce_cities.html](http://www.iihs.org/research/topics/auto_enforce_cities.html).

\(^{19}\)Massachusetts state law, Chapter 90C, section 3.

\(^{20}\)There are 350 local police agencies, one for each of the 350 towns and cities in Massachusetts (excluding Boston). There are approximately 50 different units of the state police (including various troops, headquarters, and special units), and the troops are distributed geographically throughout the state.
Table (1) shows that warnings are very prevalent in traffic stops in Massachusetts. The state police ticket about 73% of drivers they stop for speeding, and local police ticket about 40% of stopped speeders. Table (2) shows that warnings are also prevalent within specific categories of offenses other than speeding. For example, local police only ticketed about 37% of drivers who were stopped for the offense “Failure to Stop”. This “catch and release” policy is consistent with the general enforcement model where fine amounts are fixed. If traffic stops did not detect other crimes, the high frequency of warnings would be difficult to reconcile with efficient enforcement unless warnings significantly influence driving behavior. Table (3) shows that searches are very rare, and that both ticketed and warned drivers are searched. This suggests that the prevalence of warnings does not result because the police stop one group of drivers who are suspicious and another group who are not. Table (4) shows that both local and state police stop many drivers for traveling less than 10 miles-per-hour over the speed limit. The ulterior motive of detecting other crimes can explain why so many drivers are stopped for such minor violations. Warnings exist within groups of drivers who committed similar violations, which is consistent with the general enforcement model.

4 The Revenue from Tickets

Local governments in Massachusetts can feasibly take ticket revenue into account when designing their traffic enforcement policies. By state law, one-half of the revenue from traffic tickets issued by local and state police goes to the treasury of the town or city where the violation occurred.\footnote{Chapter 280, section 2. Available at http://www.mass.gov/legis/laws/mgl/280-2.htm} Table (1) showed that out-of-town drivers in Massachusetts are penalized more often after being stopped for speeding. This section uses the general enforcement framework to analyze and interpret this treatment of out-of-town drivers. After incorporating a concern for ticket revenue into the model, I show that the general enforcement framework can explain why out-of-town drivers receive any warnings, and why they are less likely to be warned. Following that, the empirical analysis tests a prediction of the model against the data.

4.1 Policy towards out-of-town drivers: Theory

In order to apply the model to this problem, I make several assumptions. The first is the most critical: I assume that the enforcement authority cannot set different probabilities of detection for in-town and out-of-town drivers. In particular, the detection probability for minor crimes, \( p(0, y) \), is the same for both groups. This assumption greatly simplifies the analysis, and it is also quite reasonable given the context. In Massachusetts, there is no numbering scheme on the license plates which identifies the town where the car is registered. Some towns have parking stickers but those are small. Therefore, it is difficult for the police to determine whether a driver is from out-of-town.
until they have already stopped the car. However, drivers from other states are easily identified
before a stop by the license plate. For this reason I exclude those drivers from the empirical analysis.

By state law, the fine amounts are also the same for in and out-of-town drivers, so I assume this
as well. I also assume that out-of-town drivers have the same distribution of benefits $G(b)$ as in-
town drivers.\(^{22}\) Finally, I assume that out-of-town drivers impose the same costs $c$ from committing
violations. This assumption is important, and in the empirical work that follows I control for the
type and degree of offense committed.

For minor crimes, the only choice variable remaining for the government is the probability of
fining a stopped driver. The police know the driver’s hometown from the driver’s license, so it is
possible for the enforcement authority to choose different rates for in-town and out-of-town drivers.
In this way, the expected fines will be different as well. I use the following notation:

\[
\pi_1(c): \text{Proportion of detected in-town crime } c \text{ criminals fined.}
\]

\[
\pi_2(c): \text{Proportion of detected out-of-town crime } c \text{ criminals fined.}
\]

\[
z_1(c) = \pi_1(c) \times p(x(c), y) \times F(c): \text{Expected fine for in-town criminals committing crime } c.
\]

\[
z_2(c) = \pi_2(c) \times p(x(c), y) \times F(c): \text{Expected fine for out-of-town criminals committing crime } c.
\]

By examining differences in warnings, only minor crimes need to be considered, because by
Proposition (1) there are no warnings for crimes which are not minor. I assume the general en-
forcement probability $p(0, y^*)$ is positive. Once that is chosen, the enforcement authority’s problem
is identical for each minor crime. It chooses the probability of fining a detected crime $c$ criminal to
maximize welfare with regards to those individuals who might potentially commit crime $c$.\(^{23}\)

For in-town drivers, the government uses the same objective function as in Proposition (2),
which takes into account both the costs and benefits of crime. With respect to out-of-town drivers,
it is logical for the government to ignore the benefits of crime because those drivers are not in the
local welfare calculation. Formally, suppose that for each minor crime $c$, the government minimized
the external costs imposed on residents by out-of-town violators, while ignoring out-of-town driver’s
benefits of crime:

\[
\max_{\pi_2(c)} -c[1 - G(z_2(c))] \quad (4)
\]

The solution to (4) is $\pi_2^*(c) = 1$. This deters the maximum number of out-of-town drivers from
committing crime $c$ for a given probability of stop and fine amount. Clearly, the data are not
consistent with this prediction because out-of-town drivers are frequently given warnings.

I now show that when the government considers fine revenue, more realistic predictions for

\(^{22}\)It will be apparent that this assumption does not affect any qualitative conclusions.

\(^{23}\)This is shown in the proof of Proposition (2). The mass of individuals who might commit minor crime $c$ is a
constant and so does not affect the solution.
the treatment of out-of-town drivers are obtained. In the general enforcement model (and the model with only one crime), the revenue from fines was treated as a welfare neutral transfer of wealth. For this reason, fine revenue did not influence the optimal policy. However, when the local government tickets an out-of-town driver, this represents a net transfer of wealth to the community. Social welfare increases because the fine came from a non-resident. Therefore, I assume that ticket revenue from out-of-town drivers adds directly to welfare, while revenue from in-town drivers remains a transfer payment. I maintain the assumption that out-of-town utility is not included in the welfare function. Under these conditions, the local government’s enforcement problem for each minor crime \( c \) with respect to out-of-town drivers is the following:

\[
\max_{\pi_2(c)} z_2(c)[1 - G(z_2(c))] - c[1 - G(z_2(c))] \tag{5}
\]

The first term in (5) is out-of-town ticket revenue, and the second term is the total cost imposed on residents by out-of-town drivers who commit crime \( c \). The solution to this problem is:

\[
\pi^*_2(c) \times p(0, y^*) \times F(c) = c + \frac{1 - G(z_2(c))}{g(z_2)} \tag{6}
\]

The \( g(\cdot) \) function is the probability distribution function for \( b \), so the expected fine for out-of-town drivers is the sum of the external cost \( c \) and the inverse hazard term \( \frac{1 - G(z_2(c))}{g(z_2)} \). Both terms are positive. A solution for \( \pi^*_2(c) \) between 0 and 1 is now possible. The analogous solution for in-town drivers is given by Proposition 2: \( \pi^*_1(c) \times p(0, y^*) \times F(c) = c \). Therefore, the model predicts that the following gap exists between the expected fines for out-of-towners and in-towners:

\[
z_2(c) - z_1(c) = \frac{1 - G(z_2)}{g(z_2)} \tag{7}
\]

This is the out-of-town penalty. Under the assumptions that the detection probability and statutory fine amounts are the same for in and out-of-town drivers, we see that:

\[
p(0, y^*) \times F(c) \times [\pi^*_2(c) - \pi^*_1(c)] = \frac{1 - G(z_2)}{g(z_2)} \tag{8}
\]

Equation (8) says the government sets a higher probability of ticketing stopped out-of-town drivers. Importantly, in the model this discrimination against out-of-town drivers is part of an efficient enforcement policy. Makowsky and Stratmann (2008) show that the Massachusetts police exercise some discretion in choosing fine amounts as well. The departures from statutory fines are small but statistically significant. Assuming that police do manipulate fine amounts to some degree, then only equation (7) would hold for certain. In practice the police could assign both higher fine amounts and lower warning probabilities to out-of-town drivers, so the empirical section addresses this.

Although the model predicts an out-of-town penalty, an empirical disparity in ticketing between
in and out-of-town drivers may not be due to a concern for revenue. The police might treat out-
of-town drivers more harshly for a number of reasons unrelated to ticket revenue. For instance, the local police might perceive a lower chance of being sanctioned for writing unjustified tickets to out-of-town drivers.

The model and the data together suggest a way to test for the revenue motive. For drivers ticketed by local police, Figure (1) shows the empirical probability that a speeding violation was faster than a given miles-per-hour. Call this probability $1 - H(mph)$, where $H(mph)$ is the cumulative distribution function for miles-per-hour over the limit. Notice that $1 - H(mph)$ is very similar for in-town and out-of-town drivers, and that above 15 miles-per-hour, $1 - H(mph)$ is concave. Assume that $H(mph)$ reflects the shape of the unknown distribution of driver’s benefits from speeding $G(b)$, so that above some $\tilde{b}$, $1 - G(b)$ is concave. If $1 - G(b)$ is concave, the following property holds:

**Monotone Hazard Rate.** The hazard rate $\frac{g(z)}{1 - G(z)}$ is increasing in the expected fine $z$.

This property holds for many distributions, including the normal, exponential, logistic, and uniform. When this holds, the inverse hazard rate $\frac{1 - G(z)}{g(z)}$ is decreasing in $z$. Under this condition, the inverse hazard term in equation (8) is predicted to be smaller for more harmful offenses which are punished with a higher expected fine. If discrimination against out-of-town drivers is motivated by revenue, the model predicts that the gap in the expected fines between in-town and out-of-town drivers should shrink for more serious violations. This prediction is the basis for the empirical test for the revenue motive which I employ below.

When the inverse hazard rate is decreasing, a marginal increase in the expected fine $z$ leads to relatively more criminals being deterred as $z$ gets larger. This means the transfer of revenue becomes less important for more serious crimes. If the police discriminate against out-of-town drivers for reasons unrelated to revenue, there is no particular reason why the out-of-town penalty should shrink for more serious violations. Since the graph of $1 - H(mph)$ was only concave above a certain value of miles-per-hour, it is reasonable to suppose that the hazard rate condition on $G(b)$ only holds above some given benefit, so incorporate this into the empirical analysis.

### 4.2 Policy towards out-of-town drivers: Empirical evidence

In this section I first show empirically that an out-of-town penalty exists. I then apply the empirical test for the revenue motive.

The out-of-town penalty could take the form of a higher chance of being ticketed, a higher fine amount charged, or some combination of the two. Therefore to determine whether an out-of-town penalty exists I estimate OLS specifications for the probability that a stopped driver receives a ticket (probit models produce similar results), and Tobit models of the fine amount charged to the stopped driver. In the Tobit specifications, I set the fine amount to zero for drivers who received

---

warnings. This is consistent with the theoretical formulation of warnings as a zero fine for breaking the law.

When estimating the out-of-town penalty, an important issue to consider is that out-of-town drivers might commit more dangerous violations. This could result in a positive estimate of the out-of-town penalty that is due to omitted variables, even when controls for the severity of the offense such as miles-per-hour over the limit are included. I address this problem in two ways. First, because most of the 350 towns in Massachusetts are geographically small, it is unlikely that in each town there are two groups of drivers, one which only drives in-town and the other which ventures out-of-town. As such, the distinction between in-town and out-of-town drivers should result only from differences in where drivers happened to be stopped.

To capture this notion, I include fixed effects for each hometown in a linear probability model for receiving a ticket after being stopped. Suppose that out-of-town drivers tend to be from certain places, such as Worcester, and that drivers from Worcester commit more dangerous offenses. The inclusion of hometown fixed effects will correct for this simple omitted variable problem by comparing how Worcester drivers are treated in Worcester versus everywhere else. As another approach, I restrict the sample to traffic stops which occurred within 10 miles of the drivers hometown. Drivers from far away might commit more dangerous offenses because they are unfamiliar with the local roads. Shrinking the sample in this way leaves out those drivers. Finally, I restrict the analysis to stops for speeding, because for other categories of stops there is no information on the severity of the offense. For speeding stops the miles-per-hour over the limit is a good indicator of offense severity (particularly when controlling for the speed limit), and it will be critical in the analysis that follows.

I include a number of control variables in each specification to account for different conditions on the roads and the severity of the offense. The controls are: Miles-per-hour over the limit, the speed limit itself, dummies for day of week and time of day (morning, afternoon, evening, predawn), and driver’s age, race (black versus white) and gender. The variable of interest is an indicator for whether the driver is from out-of-town. Column 1 in Table (5) shows the base OLS result: The local police were about 7 percentage points (standard error of 0.4 points) more likely to ticket stopped out-of-town drivers than similar in-town drivers. When including hometown fixed effects and restricting the sample to drivers stopped within 10 miles of home (Column 4), the out-of-town penalty falls to 4 percentage points (s.e. = 0.5). This is an 11.4% penalty in the chance of being ticketed, because local police ticketed 35 percent of stopped in-town speeders. A high proportion (about 72%) of stops by local police satisfied the within 10 mile restriction. The estimated out-of-town penalty for drivers stopped by the state police is only 1 percentage point and is not statistically

---

25 Distance in miles from hometown to town where stopped was calculated using the longitude and latitude of the towns as published in the U.S. Census Gazetteer (www.census.gov/cgi-bin/gazetteer).

26 The hometown fixed effects and the distance restriction do not make sense for out-of-state drivers. I already excluded those drivers from the analysis because the probability of detection can be different for them.
different from zero.

Table (6) shows the results of estimating the out-of-town penalty with OLS and Tobit specifications for the fine amount charged to the stopped driver. For the Tobits, I report the effect of being out-of-town on the expected fine amount charged (evaluated at the means of the explanatory variables). This marginal effect measures the out-of-town penalty as the combination of a higher probability of getting a ticket and a potentially higher fine amount, conditional on the fine being positive. For the base Tobit specification (1st column) the local police charged stopped out-of-town drivers 8.4 dollars (standard error of 0.49) more than similar stopped in-town drivers. This is about 7.2% of the average speeding fine charged to in-town drivers (116 dollars). Using OLS produces similar results; an out-of-town penalty of about 6.9 dollars (standard error of 0.52). On the other hand, the state police charged only an additional 3.13 dollars to out-of-town drivers. This effect falls to 1.76 dollars and is not statistically significant when the sample is restricted to drivers stopped within 10 miles of home.

The above empirical evidence suggests that an out-of-town penalty exists for drivers stopped by local police, and that the penalty does not result because out-of-town drivers commit more dangerous violations. This is the same conclusion reached by Makowsky and Stratmann (2008), and they use different empirical specifications (Heckman selection models). However, the existence of an out-of-town penalty is not sufficient to conclude that the police are motivated by ticket revenue. For example, the police might penalize out-of-town drivers because such behavior is less likely to generate political pressure on the local police department. The model is useful in this case because it offers a more subtle test of the revenue motive. Figure (2), showing the proportions of in-town and out-of-town drivers who received tickets at different values of miles-per-hour, is the most simple implementation of the empirical test for the revenue motive. It shows that the out-of-town penalty shrinks for more dangerous offenses, which is consistent with the model’s prediction for how the penalty should change if it is motivated by ticket revenue.

I now estimate several specifications to determine if the out-of-town penalty shrinks for more serious offenses when other relevant factors are controlled for. This requires estimating a triple interaction term. I do this by including the following additional explanatory variables: First, an indicator for if the violation was above 20 miles-per-hour over the limit (called $1[\text{mph} > 20]$). I used 20 miles-per-hour based on the pattern shown in Figure (2).\textsuperscript{27} Second, the interaction of $1[\text{mph} > 20]$ and the indicator for out-of-town. Third, the interaction of miles-per-hour over the limit and $1[\text{mph} > 20]$, and fourth the interaction of miles-per-hour and out-of-town. Finally, the triple interaction term of interest is out-of-town $\times$ mph $\times$ $1[\text{mph} > 20]$.

In this way, the triple interaction term captures how the effect of miles-per-hour on the outcome is different for out-of-town drivers once the miles-per-hour is above 20. If the pattern seen in Figure (2) holds up, the triple interaction term will have a negative effect on the chance of receiving a

\textsuperscript{27}The results are qualitatively similar when using a cutoff between 16 and 19 miles per hour.
ticket (and the expected fine) because the out-of-town penalty is initially positive. Estimating the marginal effect of the triple interaction is a tedious exercise in a non-linear model such as a probit, because the effect is a combination of non-linear functions of all the coefficients.\textsuperscript{28} For the reported probit and Tobit specifications, I computed the average of the marginal effects by taking discrete differences (from 0 to 1 for dummy variables, and by 1 mile-per-hour) instead of derivatives, and I estimated the standard errors of the average marginal effects by bootstrapping with 400 replications. For the OLS specifications, the marginal effect of the triple interaction term is simply its coefficient.

To get a sense of the results across different specifications, I estimated probits for the probability of being ticketed, tobits for the expected fine, and the following OLS models: for the probability of getting a ticket, for the fine amount including zeros, and for the fine conditional on it being positive. Table (7) shows the results. The marginal effect from the probit model implies that each mile-per-hour above 20 is associated with a 1 percentage point decline in the out-of-town penalty in the probability of receiving a ticket (standard error of 0.5 percentage point). The results for the linear probability model are similar. The OLS specification for the fine amount which includes zero fines shows that each mile-per-hour above 20 is associated with a reduction of 3.15 dollars (s.e.=0.97) in the out-of-town penalty. The negative coefficient on out-of-town for that specification results because most of the effect loads onto the interaction terms. The marginal effect of out-of-town for OLS (2), computed by taking the difference in the predicted fine by out-of-town status, is 7.09 dollars. OLS for only drivers who received positive fines produces a smaller coefficient on the triple interaction (-1.69 dollars), with borderline statistical significance. The Tobit model for the fine amount which includes zero fines shows that each mile-per-hour above 20 is associated with a reduction of 3.15 dollars (s.e.=0.97) in the out-of-town penalty. The negative coefficient on out-of-town for that specification results because most of the effect loads onto the interaction terms. The marginal effect of out-of-town for OLS (2), computed by taking the difference in the predicted fine by out-of-town status, is 7.09 dollars. OLS for only drivers who received positive fines produces a smaller coefficient on the triple interaction (-1.69 dollars), with borderline statistical significance. The Tobit model for the fine amount results in a still smaller marginal effect of the triple interaction on the fine (-0.26 dollars) that is not statistically different from zero. The evidence is therefore mixed in that the signs of the effects are consistent, while the magnitude and statistical significance vary across specification.\textsuperscript{29}

Table (8) shows additional specifications which either include hometown fixed effects, or restrict the sample to in-town drivers plus out-of-town drivers stopped more than 10 miles from home. Including hometown fixed effects did not change the marginal effect of the triple interaction for the linear probability model, and reduced the effect from -3.15 to -1.43 dollars in the OLS model for the fine amount including zeros. Interestingly, the size of the triple interaction effects are larger when only out-of-town drivers more than 10 miles away from home are included. Conversely, the same effects (not shown) are smaller when restricting to drivers stopped within 10 miles of home. This indicates that the evidence for a revenue motive is somewhat stronger for drivers from farther away.

Overall, the empirical evidence is suggestive that discrimination against out-of-town drivers by

\textsuperscript{28}See Cornelissen and Sonderhof (2008) for an example which provides formulas.

\textsuperscript{29}Setting a slightly different cutoff generates a similar mix of results. For example, when setting the cutoff at 17 instead of 20 mph, the effects of the triple interaction term for the three fine amount specifications in Table (7) are as follows: Tobit -0.02 (s.e.=0.29), OLS (2) -1.72 (s.e.=0.64), OLS (3) -1.39 (s.e.=0.65).
the local police is motivated by ticket revenue. If the police simply gave in-town drivers preferential treatment, the out-of-town penalty would be the same across different degrees of violations. Instead, across different specifications and sample restrictions, the out-of-town penalty in the chance of being ticketed and the expected fine shrinks for more serious violations. The evidence for the revenue motive is weakest in the Tobit models, where the relevant marginal effects were all negative but were smaller and not statistically different from zero.

5 Conclusion

This paper showed that under certain conditions, releasing criminals with only a warning instead of requiring them to pay a fine can be an aspect of an efficient enforcement policy. The result holds even if the behavior of potential criminals is unaffected by warnings. The required conditions are that the enforcement technology detects other crimes, and that fine amounts are either fixed or cannot be finely adjusted. Traffic enforcement by police working for local governments satisfies these conditions, so the widespread practice of giving warnings to drivers stopped for breaking traffic laws can in theory be efficient even if drivers don’t respond to warnings.

Empirically, stopped out-of-town drivers in Massachusetts were more likely to receive a ticket than similar in-town drivers. I used an empirical test suggested by the model to determine if this discrimination is motivated by ticket revenue. The balance of the evidence indicates that the revenue motive is at work, because the penalty imposed on out-of-town drivers tends to disappear for more serious violations. Using different methods, Makowsky and Stratmann (2008) also concluded that a revenue motive for ticketing exists in Massachusetts, as did Garrett and Wagner (2008) for North Carolina. I showed that this particular revenue motive does not conflict with effective law enforcement, if one thinks of the revenue from out-of-town traffic tickets as a net transfer of wealth to the local community.
6 Appendix

Assume that $y^* > 0$, so that all crimes are detected with at least probability $p(0, y^*)$.

Proof of Proposition 1

Suppose that $\pi^*(c) < 1$ and $x^*(c) > 0$. Then reduce $x^*(c)$ and increase $\pi^*(c)$ to hold the expected fine $z(c) = \pi^*(c) \times p(x^*(c), y^*) \times F(c)$ constant. Then the same number of individuals commit crime $c$ but fewer resources are used, a contradiction.

Proof of Proposition 2

For minor crimes where $x^*(c) = 0$, the government takes as given $p(0, y^*)$ and $F(c)$. Therefore for a minor crime $c$, the government’s problem is:

$$\max_{\pi(c)} \int_{\pi(c) \times p(0, y^*) \times F(c)}^{\infty} (b - c) dG(b)s(c)$$

The mass of individuals who might commit crime $c$ is $s(c)$. Since $s(c)$ is just a constant it does not affect the solution and therefore can be dropped. Where $z = \pi(c) \times p(0, y^*) \times F(c)$, the first order condition is:

$$- [\pi^*(c)p(0, y^*)F(c) - c] g(z) d\pi(c)p(0, y^*)F(c) = 0$$

This simplifies to $\pi^*(c) \times p(0, y^*) \times F(c) = c$. Suppose $\pi^*(c) = 1$ and $p(0, y^*) \times F(c) > c$. Then at no cost, social welfare can be increased by reducing $\pi^*(c)$ so that an individual whose benefit $b$ exceeds the social cost $c$ will be induced to commit the crime. Because all $c > 0$, $\pi^*(c) > 0$. 
### 7 Tables and Figures

Table 1: Proportion of traffic stops resulting in a ticket.

<table>
<thead>
<tr>
<th></th>
<th>Proportion Ticketed</th>
<th>Number of stops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Other offenses</td>
<td>Speeding</td>
</tr>
<tr>
<td><strong>State Police</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.55</td>
<td>0.75</td>
</tr>
<tr>
<td>In-town</td>
<td>0.54</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>Local Police</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.46</td>
<td>0.43</td>
</tr>
<tr>
<td>In-town</td>
<td>0.46</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Boston Police</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>In-town</td>
<td>0.48</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Excludes out-of-state drivers.

Table 2: Proportion of stops for selected other offenses resulting in a ticket.

<table>
<thead>
<tr>
<th></th>
<th>Proportion Ticketed</th>
<th>Number of stops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Other offenses</td>
<td>Speeding</td>
</tr>
<tr>
<td><strong>State Police</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.56</td>
<td>0.84</td>
</tr>
<tr>
<td>In-town</td>
<td>0.57</td>
<td>0.84</td>
</tr>
<tr>
<td><strong>Local Police</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.37</td>
<td>0.86</td>
</tr>
<tr>
<td>In-town</td>
<td>0.37</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Boston Police</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.44</td>
<td>0.58</td>
</tr>
<tr>
<td>In-town</td>
<td>0.42</td>
<td>0.55</td>
</tr>
</tbody>
</table>

(1) Failure to Stop (2) No Inspection Sticker (3) Seat Belt Violation
Excludes out-of-state drivers.
Table 3: Number of searches.

<table>
<thead>
<tr>
<th></th>
<th>Warned Searches</th>
<th>Warned Total Stops</th>
<th>Ticketed Searches</th>
<th>Ticketed Total Stops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Police</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Offenses</td>
<td>80</td>
<td>20,545</td>
<td>181</td>
<td>17,357</td>
</tr>
<tr>
<td>Speeding</td>
<td>44</td>
<td>30,279</td>
<td>76</td>
<td>21,335</td>
</tr>
<tr>
<td>State Police</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Offenses</td>
<td>57</td>
<td>5,835</td>
<td>92</td>
<td>6,969</td>
</tr>
<tr>
<td>Speeding</td>
<td>12</td>
<td>5,512</td>
<td>89</td>
<td>17,440</td>
</tr>
</tbody>
</table>

Excludes out-of-state drivers.

Table 4: Proportion ticketed by MPH over speed limit.

<table>
<thead>
<tr>
<th></th>
<th>Proportion Ticketed</th>
<th>Number of stops</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Out-of-town</td>
<td>In-town</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPH over limit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-10</td>
<td>0.40</td>
<td>0.31</td>
<td>7,328</td>
<td>3,372</td>
</tr>
<tr>
<td>11-15</td>
<td>0.28</td>
<td>0.21</td>
<td>15,064</td>
<td>7,571</td>
</tr>
<tr>
<td>16-20</td>
<td>0.55</td>
<td>0.48</td>
<td>9,945</td>
<td>4,326</td>
</tr>
<tr>
<td>21-25</td>
<td>0.77</td>
<td>0.70</td>
<td>2,631</td>
<td>1,151</td>
</tr>
<tr>
<td>26 +</td>
<td>0.91</td>
<td>0.89</td>
<td>856</td>
<td>389</td>
</tr>
<tr>
<td>Local Police</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-10</td>
<td>0.54</td>
<td>0.49</td>
<td>5,905</td>
<td>795</td>
</tr>
<tr>
<td>11-15</td>
<td>0.77</td>
<td>0.61</td>
<td>3,870</td>
<td>443</td>
</tr>
<tr>
<td>16-20</td>
<td>0.92</td>
<td>0.87</td>
<td>3,978</td>
<td>423</td>
</tr>
<tr>
<td>21-25</td>
<td>0.97</td>
<td>0.96</td>
<td>2,336</td>
<td>246</td>
</tr>
<tr>
<td>26 +</td>
<td>0.98</td>
<td>0.97</td>
<td>1,229</td>
<td>197</td>
</tr>
<tr>
<td>State Police</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Excludes out-of-state drivers.
Figure 1: Probability that violation was faster than MPH (local police).

Table 5: Out-of-town penalty in the probability of being ticketed for speeding (OLS).

<table>
<thead>
<tr>
<th>Ticketed (Yes=1)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out-of-town</td>
<td>0.07**</td>
<td>0.07**</td>
<td>0.04**</td>
<td>0.04**</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Linear mph</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hometown fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>10 miles from home</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Driver demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Day and time</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>52,398</td>
<td>37,394</td>
<td>52,398</td>
<td>37,394</td>
<td>19,314</td>
</tr>
</tbody>
</table>

Dependent variable: Ticketed after stop (Yes=1, No=0)
Heteroskedastic robust standard errors, **p<0.05
Table 6: Out-of-town penalty in fine amount charged.

<table>
<thead>
<tr>
<th>Fine Amount (warning=0)</th>
<th>Tobit</th>
<th>Tobit</th>
<th>OLS</th>
<th>OLS</th>
<th>Tobit</th>
<th>Tobit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out-of-town</td>
<td>8.40**</td>
<td>8.25**</td>
<td>6.91**</td>
<td>6.96**</td>
<td>3.13**</td>
<td>1.76</td>
</tr>
<tr>
<td>(effect on $E[fine</td>
<td>X]$)</td>
<td>(0.499)</td>
<td>(0.553)</td>
<td>(0.517)</td>
<td>(0.570)</td>
<td>(1.56)</td>
</tr>
<tr>
<td>Linear mph</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hometown fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>10 miles from home</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Driver demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Day and time</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>46,875</td>
<td>33,889</td>
<td>46,875</td>
<td>33,889</td>
<td>16,736</td>
<td>5,921</td>
</tr>
</tbody>
</table>

Dependent variable: Fine amount charged (warning=0).
Tobit: Marginal effect of out-of-town on fine amount, at means. **$p<0.05$

Figure 2: Proportion ticketed by MPH over speed limit (local police).
Table 7: Shrinking out-of-town penalty in speeding tickets.

<table>
<thead>
<tr>
<th>Local police only</th>
<th>Prob(Ticketed=1)</th>
<th>Fine amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>Probit</td>
</tr>
<tr>
<td>Out-of-town×mph×1[mph &gt; 20]</td>
<td>−0.005∗ (0.0028)</td>
<td>−0.01∗ (0.005)</td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.08** (0.021)</td>
<td>0.07** (0.004)</td>
</tr>
<tr>
<td>Miles-per-hour</td>
<td>0.03** (0.001)</td>
<td>0.03** (0.001)</td>
</tr>
<tr>
<td>1[mph &gt; 20]</td>
<td>0.38** (0.054)</td>
<td>0.16** (0.030)</td>
</tr>
<tr>
<td>10 miles from home</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Driver demographics</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Day and time</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>52,398</td>
<td>52,398</td>
</tr>
</tbody>
</table>

Probit: Average marginal effects on the probability of being ticketed.
Tobit: Average marginal effects on the fine amount.
Probit and Tobit standard errors are bootstrapped with 400 replications. **p<0.05.
OLS (3) includes only observations with positive fines.
Table 8: Shrinking out-of-town penalty, additional specifications.

<table>
<thead>
<tr>
<th></th>
<th>Local police only</th>
<th>Prob(Ticketed=1)</th>
<th>Fine amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OLS (1)</td>
<td>Probit</td>
</tr>
<tr>
<td>Out-of-town×mph×1[mph &gt; 20]</td>
<td>−0.005</td>
<td>−0.016**</td>
<td>−0.61</td>
</tr>
<tr>
<td></td>
<td>(0.0029)</td>
<td>(0.007)</td>
<td>(0.406)</td>
</tr>
<tr>
<td>Out-of-town</td>
<td>0.05**</td>
<td>0.07**</td>
<td>8.54**</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.005)</td>
<td>(0.678)</td>
</tr>
<tr>
<td>Miles-per-hour</td>
<td>0.03**</td>
<td>0.03**</td>
<td>4.14**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>1[mph &gt; 20]</td>
<td>0.42**</td>
<td>0.15**</td>
<td>20.84**</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.033)</td>
<td>(2.86)</td>
</tr>
<tr>
<td>Hometown fixed effects</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>More than 10 miles</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Driver demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Day and time</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>52,398</td>
<td>31,189</td>
<td>28,052</td>
</tr>
</tbody>
</table>

Probit: Average marginal effects on the probability of being ticketed.
Tobit: Average marginal effects on the fine amount.
Probit and Tobit standard errors are bootstrapped with 400 replications.**p<0.05

8 References

Durose, Matthew; Erica Smith; Patrick Langan. “Contacts between Police and the Public, 2005”.

26
Farrell, Amy; Dean McDevitt; Lisa Bailey; Carsten Andresen; Erica Pierce. “Massachusetts Racial and Gender Profiling Final Report: Executive Summary”, Northeastern University Institute on Race and Justice, May 2004.