#### **Contributions**

Kong-Pin Chen and Tsung-Sheng Tsai\*

# Judicial Torture as a Screening Device

**Abstract:** Judicial torture to extract information or to elicit a confession was a common practice in pre-modern societies, both in the east and the west. This paper proposes a positive theory for judicial torture. It is shown that torture reflects the magistrate's attempt to balance type I and type II errors in the decision-making, by forcing the guilty to confess with a higher probability than the innocent, and thereby decreases the type I error at the cost of the type II error. Moreover, there is a non-monotonic relationship between the superiority of torture and the informativeness of investigation: when investigation is relatively uninformative, an improvement in technology used in the investigation actually lends an advantage to torture so that torture is even more attractive to the magistrates; however, when technological progress reaches a certain threshold, the advantage of torture is weakened, so that a judicial system based on torture becomes inferior to one based on evidence. This result can explain the historical development of the judicial system.

**Keywords:** torture, type I and type II errors, evidence

DOI 10.1515/bejte-2014-0023

### 1 Introduction

Judicial torture for the purpose of eliciting information was a common practice in pre-modern societies. In the west, it emerged in Greek law and continued in Roman law. Its history can then be traced through the Middle Ages, down to the legal reforms of the eighteenth century and the abolition of torture in criminal legal procedure in the nineteenth century in most parts of Europe (Peters 1985, 5). In China, judicial torture formed an important part of the imperial legal codes

<sup>\*</sup>Corresponding author: Tsung-Sheng Tsai, Department of Economics, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan, E-mail: tstsai@ntu.edu.tw Kong-Pin Chen, Research Center for Humanities and Social Sciences, Academia Sinica, No. 128, Academia Road, Nangang, Taipei 11529, Taiwan, E-mail: kongpin@gate.sinica.edu.tw

from the first Empire (Qin dynasty, 221–207 BC) to the last Qing dynasty (1644–1911) (Shen 1985). Various forms of judicial torture were also widely implemented in Muslim, African and various Asian societies before our time (Lea 1971, 203–8). It is indeed hard to find any pre-modern society that did not rely on torture for the gathering of evidence in judicial proceedings.

The modern man, for whom judicial torture is not only immoral but also irrational, may imagine that it was inflicted exclusively on the accused to obtain a confession. Nothing is further from the truth. First, judicial torture was officially institutionalized, as its practice was explicitly written into many legal codes. Second, torture was explicitly applied not only to the suspect, but also to the witness, and even to the plaintiff. Witnesses were tortured in Greek and Roman cases, and this continued in the medieval period and until the late eighteenth century (Peters 1985, 18, 69). The historical development of judicial torture has indicated that it was an institutionalized practice, rather than a manifestation of a magistrate's abuse of power.

During the Middle Ages, European courts had increasingly used torture as a legitimate means to extract confessions or to obtain the accomplice's identity or other information regarding a crime. In 1252, Pope Innocent IV issued a papal bull which authorized the use of torture by inquisitors. The application of torture began to appear in significant numbers from around the thirteenth century onward and peaked in the sixteenth century. However, starting from around the mid-eighteenth century, the practice of judicial torture in the west was gradually replaced by a system which was based on evidence. Its abolition was a long and gradual process that lasted from the mid-eighteenth to the early nineteenth centuries. The conventional historical explanation of this abolition movement relies heavily on the influence of the humanists of the Enlightenment. However, experts on jurisprudence usually find this explanation too loose, and prefer to explain the phenomenon by changes in the judicial system itself. According to John Langbein, a famous legal scholar, the abolition was largely due to the emergence of the new law of proof. During this period, the courts gradually imposed new and less rigorous punishments according to a less strict standard of proof, one of persuasion rather than certainty. Since certainty was no longer the only requirement to put the accused in prison, torture became unnecessary. "Only when confession evidence was no longer necessary to convict the guilty could European law escape its centuries of dependence on judicial torture" (Langbein 1983, 1555-6). This explanation, however, is incomplete because, as noted by Damaška (1978) and Silverman (2001), torture continued to be used well after the change in the law of proof.

In this paper, we propose a positive model for judicial torture based on the theory of information economics to explain its rise, decline and eventual abolishment in the legal system. We explain why under a certain environment a system based on torture might become a prevailing judicial system. Besides the pain inflicted on the accused, there are two consequences of torture. On the one hand, since some innocent suspects are tortured to confession, torture increases the chance that an innocent suspect will be wrongfully convicted (i.e. a type I error). On the other hand, if investigation is of informational value, in the sense that it is more likely to find evidence against the guilty suspects during investigation, then (if the magistrate uses torture as a threat when such evidence is found) the expected cost of denying a crime will be greater for the guilty than the innocent. This reduces the chance that a guilty suspect will be wrongfully released (i.e. a type II error). The goal of the magistrate is to balance the costs of type I and type II errors, as well as the cost of torture such as the pain imposed on the suspect or the witness.

The main argument of this paper is that the attractiveness of torture is crucially determined by the precision of investigation. It will be shown that if little information is revealed during the investigation, the gain in reducing a type II error (by torture) will outweigh the cost of the increase in a type I error. In such a case, torture is the "optimal" choice for the magistrates. However, there is a non-monotonic relationship between the superiority of torture and the precision of investigation: when investigation is relatively uninformative, an improvement in technology used in the investigation can lend an advantage to torture so that torture is even more attractive to the magistrates; however, when technological progress reaches a certain threshold, the advantage of torture is weakened, so that a system solely based on evidence will overtake torture as a better system.

Based on our theory, the rise and decline of judicial torture in European history can be explained in a more satisfactory way if we look at another factor other than the humanitarian considerations or the change in the law of proof: the employment of medical science in law. Between the thirteenth and the sixteenth centuries, the trend of the application of medical knowledge in judicial cases exhibited a stable increase, while at about the same time torture was used more and more intensively and its application reached its peak in the sixteenth century. However, later developments in medical science between the seventeenth and nineteenth centuries, including fingerprinting, physical matching and the use of precise measurements of the human body structure (anthropometry), indicates that, with those technological breakthroughs, the precision of investigation had passed that threshold. Past human activity could now be specified without relying exclusively on the confessions of the guilty or witnesses, often obtained by torture in earlier times (Parker 1983, 432-3; Fullmer 1980, 27). During this period of time, the number of incidents of judicial torture

began to decrease substantially in many European regions. The application of modern scientific technology had apparently reached a mature stage in the late nineteenth century when Hans Gross published the first classic in the specialty, Handbuch für Untersuchungsrichter als System der Kriminalistik, in 1893 (published in English in 1907 under the title Criminal Investigation). Thereafter, a system based on evidence finally dominated, and judicial torture was eventually abolished from the statutory law by the end of the nineteenth century in most European countries.

Our model can also explain why it may be optimal for the magistrate to torture the witness or the plaintiff. If the witness or the plaintiff is also tortured, then only those witnesses or plaintiffs who are more sure of the suspect's crime will come to the court. Thus, the average quality of the cases entering the court will increase. This implies a lower type I error when the magistrate also tortures the suspect in the legal proceedings. If this reduction in the type I error is sufficiently large, then torturing the plaintiff or the witness is indeed the magistrate's optimal choice. Note that this explanation does not rely on the possibility that the magistrate might be sadistic. On the contrary, this is true even if the torturee's pain enters negatively into the magistrate's utility function.

The insight that resources need to be invested in order to overcome informational barriers is a common theme in the theoretical economics literature. For example, workers take costly actions (mainly education) to signal their abilities (e.g. Spence 1973); firms engage in costly predation to signal their toughness against new entrants (e.g. Kreps and Wilson 1982; Milgrom and Roberts 1982) and bargainers use costly delays to signal their time-patience during bargaining (e.g. Rubinstein 1985). Our theory is related to the literature in which a firm might profitably utilize a strategy that raises the cost of all competing firms in the market (Salop 1979). In our model the magistrate uses costly torture to partially resolve the uncertainty regarding the type of the accused.

There is surprisingly scant formal modeling of torture. Wantchekon and Healy (1999) analyze torture as a game with incomplete information between the state (torturer) and the victim. The latter does not know whether the torturer is sadistic (deriving pleasure from torture) or professional (incurring cost when torturing); the former is uncertain of whether the victim is strong or weak. The victim possesses a certain information, and the state decides whether or not to torture to extract that information. They show that in equilibrium there will be a positive probability of torture, even for the professional torturer, as he tries to test the victim's type. In our model, the magistrate is never sadistic, as he also suffers from torture. Therefore, our model does not explain torture in terms of some torturer being sadistic. Mialon et al. (2012) build up a theoretical model to analyze the effect when a government agency is allowed to torture the suspects

when evidence of terrorist involvement is high. They show that allowing torture reduces the agency's effort to counter terrorism by means other than torture. This reduces the quality of the agency's information in deciding whether to torture, which in turn prompts them to torture even in cases where evidence is low. Their paper differs from ours in that they are mainly concerned with the possibility of a "slippery slope" in which the agency might be tempted (albeit optimally) to torture the suspects with low evidence when they are only allowed to do so in high evidence cases. In Chen et al. (2009), it is shown that even if torture is blind in distinguishing the innocent and the guilty (i.e. they have the same probability of confession under torture), as long as there is two-sided informational asymmetry regarding the degree of the magistrate's willingness to torture and the defendant's willingness to endure, torture can be an equilibrium outcome.1

The remainder of this article is organized as follows. Section 2 introduces the model. Section 3 discusses the equilibrium outcomes in the evidence-based system and Section 4 analyzes those in the game of torture. Section 5 compares the social welfare under these two systems and discusses how the optimality is affected by the informativeness of investigation. Section 6 extends the basic model to consider the situation where the magistrate can torture the plaintiff or witness. Section 7 reviews the historical development of torture and Section 8 concludes the paper. The details of the proofs are relegated to the Appendices.

### 2 The model

Suppose that an accused suspect in a certain crime case is brought to the court ruled by a magistrate. The suspect is either guilty ( $\theta = G$ ) or innocent ( $\theta = I$ ), which is his private information. The magistrate needs to reach a verdict on whether he is guilty or not. The prior belief of the magistrate that the suspect is guilty is q, with 0 < q < 1. That is,  $Prob(\theta = G) = q^2$ .

We consider two possible types of judicial system: the evidence-based system and the torture system. In the evidence-based system, the magistrate

<sup>1</sup> Other papers that utilize game-theoretical models to investigate how rules of proceeding in criminal cases affect information transmission are Seidmann (2005), Seidmann and Stein (2000) and Sanchirico (2000, 2001). In particular, Sanchirico (2000), like our paper, also tries to explain a certain aspect of legal history with a theoretical framework.

<sup>2</sup> In Section 6 of the paper, we will investigate how the magistrate can screen the cases by torturing the plaintiffs and the witnesses. As a result, the value of q can be influenced by the decision regarding whether to bring the case to court.

conducts an investigation, which can improve the precision of his information but not resolve the uncertainty he faces. Specifically, the magistrate draws a signal (evidence) from the set  $S = \{g, i\}$ . If the suspect is indeed guilty, then with probability  $q_G$  (resp.  $1-q_G$ ) the magistrate will draw a piece of evidence s=g(resp. s = i). If the suspect is innocent, then with probability  $q_I$  (resp.  $1 - q_I$ ) the magistrate draws g (resp. i). We assume that  $1 > q_G > q_I > 0$ , i.e. the investigation is informative in the sense that the magistrate is more likely to draw g when he faces a guilty defendant than when he faces an innocent defendant. Let  $\hat{q}$  be the posterior that the defendant is guilty. The magistrate then decides whether or not to convict the suspect based on the posterior. If the defendant is convicted, he is subject to a legal penalty causing a disutility P. Otherwise, he is released, in which case his utility is 0.

Since investigation is imperfect, there are two possible errors that the magistrate can commit. First, he might wrongfully convict a defendant who is actually innocent. We call this a type I error. Second, he might wrongfully release a defendant who is actually guilty. We call this a type II error. Assume that a type I error causes a loss of  $L_1$ , and that a type II error causes a loss of  $L_2$ , to the magistrate's utility.<sup>3</sup> The objective of the magistrate is to minimize the value of his expected loss, which is  $\hat{q}L_2$  if he releases the accused and  $(1-\hat{q})L_1$  if he convicts him, by choosing whether to convict the suspect after drawing the evidence.

In the torture system, a confession from the defendant is required for a conviction. In other words, the magistrate cannot convict the suspect if the suspect does not confess. For that purpose the magistrate is allowed to torture.<sup>4</sup> The defendant is given an option of whether to confess or not. If he does, the magistrate will not conduct any investigation but will convict him directly. In this case, he is subject to a legal penalty for the crime, which causes a disutility of P. If he does not, the magistrate will go through the same investigation as in the evidence-based system. Based on the evidence drawn, he will decide whether or not to torture the defendant. If he does not, then the defendant will be released, which results in a utility of 0 for the latter. If he tortures, let T be the defendant's disutility from being tortured. In this case, his total expected

**<sup>3</sup>** Although  $L_1$  and  $L_2$  can also be taken to be a certain measure of the "social welfare" loss, for our purpose it suffices that they represent the subjective loss of the magistrate or the judicial system designer.

<sup>4</sup> That is, a confession is sufficient to found a conviction. In early English common-law trials, confessions were always allowed as evidence, even though torture was also used to elicit them. It was not until the mid-eighteenth century that judges in England began to admit only confessions that they deemed trustworthy. See "Confession," West's Encyclopedia of American Law (2005).

disutility will be T + rP, where r is the probability that the defendant will confess when tortured. This setup can be endogenized by viewing torture as a war-ofattrition game between the magistrate and the defendant. If r is the probability that the defendant will concede before the magistrate (in the war of attrition), and T is the expected length it takes for one of the sides to concede, then the expected utility of the defendant, when he is tortured, is exactly T + rP.

The magistrate's (dis)utility is assumed to be  $p_1L_1 + p_2L_2 + kT$ , where  $p_1$  and  $p_2$  are the probabilities of committing type I and II errors, respectively. Note that when a level of torture *T* is imposed on the accused, it also causes a loss of *kT* in the magistrate's utility.<sup>6</sup> Therefore, unlike the simple explanation that the magistrate tortures simply because he is sadistic, in our model the magistrate is known by the defendant to dislike torture. The magistrate is benevolent in that his aim is to minimize both the sum of the expected losses caused by the two types of errors and, if he tortures, the defendant's pain due to torture. Since the magistrate's decision will affect the defendant's confession decision and vice versa, this is a game played between the magistrate and the defendant. The torture system is substantially more complicated than the evidence-based system, as it involves the strategic interaction of the two parties.

Before proceeding with the analysis, there are several points regarding our setup that are worth mentioning. First of all, the assumption  $q_I < q_G$  is crucial to our result that the torture system can sometimes be superior to the evidencebased system. Since it is more likely that a guilty evidence will be found when the defendant is guilty, the expected cost of denying a crime will be greater for the guilty than the innocent if the magistrate ever applies torture. Thus, it will be more likely that the guilty will confess when facing a threat of being tortured. By using this "screening device," the magistrate can be more certain about the defendant's type during the investigation if he does not confess. This will decrease the chances of a guilty defendant being wrongfully released so that the type II error will be reduced. Such an advantage in information elicitation brought about by torture is the main reason for its superiority.

Secondly, investigation takes place after the initial questioning but before torture. Given this, the optimal strategy of a suspect, if he intends to confess, is actually to deny his crime initially, and to confess immediately at precisely the moment when the magistrate draws a guilty signal and decides to torture. Our specification assumes that the magistrate can make a credible commitment in

<sup>5</sup> For a formal model of torture as a war-of-attrition, see Chen et al. (2009).

**<sup>6</sup>** We assume that k is the same for both types. It is possible that the magistrate is more concerned about the pain suffered by the innocent, that is,  $k_I > k_G$ ; however, this setup will lead to a similar result to the situation where the magistrate is more concerned about  $L_1$ .

that if he decides to torture, he commits to torture by an amount T, regardless of whether the suspect confesses before or during torture. This assumption of commitment greatly simplifies the model, as there are only two consequences for a suspect who does not confess: either he is released, or he is tortured and confesses. The tradeoff facing the suspect is indeed simple: if he confesses before investigation, he suffers a disutility P; while if he does not confess, he suffers an extra pain of T but may be released with some probability. This specification is equivalent to the real world plea bargaining, in which the suspect's penalty is reduced (by an amount of T in our specification) if he chooses to confess before investigation.

Finally, we assume that the innocent and the guilty defendants have the same probability of confessing under torture (i.e. r is independent of the type  $\theta$ ). Historical evidence suggests that torture is expected to function in such a way that the accused defendant will reveal details of his crime which no innocent person possesses (and be convicted).<sup>8</sup> If the system functions perfectly, then only guilty defendants will confess under torture, as no innocent defendant can offer evidence of the crime. The assumption that torture is blind in distinguishing between the innocent and the guilty actually places a heavier burden on our proof. If it is assumed that the guilty is more likely to confess under torture (i.e.  $r_G > r_I$ ), our argument will hold even more generally. That is, we will actually show that a system based on torture might incur a lower social loss than one based on evidence, even if torture has no power in distinguishing between the innocent and the guilty.

To sum up, the events under the evidence-based system proceed in the following order: (1) Nature determines whether the suspect is guilty or innocent (i.e. whether  $\theta = G$  or  $\theta = I$ ). (2) The magistrate draws on evidence. (3) Based on the evidence, the magistrate decides whether to convict or release the accused. The events under the torture system proceed in the following order: (1) Nature determines whether the suspect is guilty. (2) The suspect decides whether to confess or not. If he confesses, then he is convicted and no further investigation is conducted. In this case, he is subject to a legal penalty of P. If he does not confess, then (3) the magistrate conducts an investigation by drawing on evidence. Based on the evidence, he decides whether to torture the suspect or not.

<sup>7</sup> Langbein (1978) states that "there are remarkable parallels in origin, in function, and even in specific points of doctrine, between the law of torture and the law of plea bargaining." It was often the case that a suspect who had confessed under torture recanted when he was asked to confirm his confession. The magistrate then had to commit to torturing the suspect repeatedly until the factual details of the crime were elicited. The suspect would quickly learn that it was better to "voluntarily" confess the crime to save himself from being tortured further. Therefore, the practice of torture closely resembles plea bargaining.

<sup>8</sup> See Langbein (1974, 1977, 1978, 1983).

If not, the suspect is released. (4) If the magistrate tortures, then after suffering a disutility of torture T, the defendant confesses with probability r.

## 3 The evidence-based system

Under the evidence-based system, the magistrate draws on evidence, and decides whether to convict the suspect solely based on the posterior derived from the evidence. If he draws a piece of evidence i, his posterior that the suspect is guilty will be

$$q_i' = \frac{q(1-q_G)}{q(1-q_G) + (1-q)(1-q_I)}$$
.

Similarly, if he draws g, the posterior will be

$$q_g' \equiv \frac{qq_G}{qq_G + (1-q)q_I}$$
.

Thus, if the evidence collected is  $s \in \{i, g\}$ , then the expected loss will be  $(1-q_s')L_1$  if the magistrate convicts the suspect and is  $q_s'L_2$  if he releases the suspect. Whether he should convict the suspect or not depends on the evidence he draws and the relative sizes of  $(1-q_s')L_1$  and  $q_s'L_2$ . To make our discussion non-trivial, we make the following assumption:

**A1.** 
$$\frac{q_G}{q_I}L_2 > \frac{1-q}{q}L_1 > \frac{1-q_G}{1-q_I}L_2$$
.

If A1 does not hold, then regardless of the evidence drawn, either he never convicts (if the second inequality fails to hold) or he never releases a suspect (if the first inequality fails to hold). Investigation thus does not have any function. If A1 holds, then it can be easily checked that  $q'_gL_2 > (1-q'_g)L_1$  and  $q_i'L_2 < (1-q_i')L_1$ . Thus, the magistrate will convict (release) the suspect if he draws g(i). We therefore have the following lemma:

**Lemma 1.** Suppose that A1 holds. Then, under the evidence-based system, the magistrate convicts the suspect if and only if he draws g.

## 4 The game of torture

As a benchmark, we first consider the case where no investigation is conducted. Since the prior belief of the magistrate that the suspect is guilty is q, if the magistrate decides to torture the suspect, the probability of making a type I error is (1-q)r when the suspect is tortured to confess, and the probability of making a type II error is q(1-r) when he still denies the crime even after being tortured. On the other hand, the probability of making a type II error is q if the suspect is released. Thus, without an investigation, the magistrate's expected loss when a suspect is tortured is  $r(1-q)L_1 + (1-r)qL_2 + kT$  and is  $qL_2$  if the defendant is released. We make the following assumption:

**A2.** 
$$\frac{1-q}{q}L_1 + \frac{kT}{rq} > L_2 > \frac{kT}{r}$$
.

The first inequality means that under the prior belief q, the expected social loss from torture is greater than that from releasing the suspect. This implies that the magistrate will not torture when no investigation is conducted. If this assumption is not satisfied, since an investigation can only reduce the expected cost under torture, the magistrate will torture the suspect anyway regardless of what signal he receives during investigation, which is, however, an uninteresting case The second inequality is obtained by assuming  $r(1-q)L_1 + (1-r)qL_2 + kT < qL_2$  when q = 1, which means that if the magistrate believes that the suspect is guilty with probability 1, the loss from the type II error if he is released is always greater than that from torture, so that the magistrate may want to torture the suspect. If this condition is violated, the magistrate will never torture even if he believes that the suspect is guilty with probability 1. In this case, there will never be torture at all, and no suspect will confess. This is again a less interesting situation because torture plays no role.

### 4.1 The optimal decision of the magistrate

Suppose that after the defendant has denied the crime, and the evidence is collected, the magistrate is considering whether to torture. Note that since the defendant's type is his private information, his decision will be based on whether he is innocent or guilty, and this decision will in turn play a role in determining the posterior of the magistrate. Assume that the innocent defendant denies the crime with probability  $v_I$ , and the guilty denies the crime with probability  $v_G$ . Then the posterior of the magistrate that the suspect is guilty, when he draws g, is

$$\hat{q}(\theta = G|v_G, v_I, g) = \frac{qv_Gq_G}{qv_Gq_G + (1 - q)v_Iq_I} \equiv \hat{q}_g.$$

On the other hand, the magistrate's posterior, when he draws i, is

$$\hat{q}(\theta = G|\nu_G, \nu_I, i) = \frac{q\nu_G(1 - q_G)}{q\nu_G(1 - q_G) + (1 - q)\nu_I(1 - q_I)} \equiv \hat{q}_i.$$

It can be easily seen that  $\hat{q}_g > \hat{q}_i$ . Based on this posterior, the expected loss for the magistrate is  $r(1-\hat{q}_s)L_1+(1-r)\hat{q}_sL_2+kT$  if he tortures the suspect, where s=i,g. If he decides not to torture and releases the suspect, the expected loss is  $\hat{q}_s L_2$ . Therefore, the magistrate tortures the suspect if and only if

$$(1-\hat{q}_s)L_1 + \frac{kT}{r} \le \hat{q}_s L_2. \tag{1}$$

### 4.2 Optimal decision of the suspect

The decision of a suspect of type  $\theta$  is to choose the value of  $v_{\theta}$  to maximize his expected utility, which is

$$-v_{\theta}(q_{\theta}x_{g}+(1-q_{\theta})x_{i})(T+rP)-(1-v_{\theta})P,$$

where  $x_g$  and  $x_i$  are the probabilities that the magistrate will torture after he draws g and i, respectively. Since  $\hat{q}_g > \hat{q}_i$ , according to eq. [1], it is obvious that  $x_g \ge x_i$  in equilibrium. The following two lemmas are very useful for solving the equilibria later.

**Lemma 2.** There does not exist a separating equilibrium. That is, there does not exist an equilibrium in which  $v_G = 0$  and  $v_I = 1$ .

*Proof.* Consider a separating equilibrium in which  $v_I = 1$  and  $v_G = 0$ . Then by Bayes' rule, it is easy to see that the posterior for the magistrate is  $\hat{q}_i = \hat{q}_g = 0$ . That is, he believes that a denying defendant is innocent, regardless of the evidence drawn. Then according to eq. [1], he will not torture the denying defendant (i.e.  $x_i = x_g = 0$ ). However, if this is the case, then the guilty defendant will surely deviate from  $v_G = 0$ .

Lemma 2 implies that, in any equilibrium, the pool of denying defendants (if there are any) will comprise both innocent and guilty suspects. An important consequence of this is that both types of denying defendant will face the same posterior of the magistrate, and will be tortured with the same probability. That is, as we have already mentioned,  $x_g$  and  $x_i$  are independent of  $\theta$ .

**Lemma 3.** In any equilibrium, the guilty defendant will confess with a higher probability than the innocent one; i.e.  $v_I \ge v_G$ .

*Proof.* If a defendant confesses before being tortured, then he is subject to a legal penalty P, regardless of his type. If he denies, then his expected utility is  $-[q_Gx_g+(1-q_G)x_i](T+rP)$  if he is guilty, and  $-[q_Ix_g+(1-q_I)x_i](T+rP)$  if he is innocent. Since  $q_G > q_I$  and  $x_g \ge x_i$ , the consequence of denying is more serious for the guilty defendant. As a result, it must be the case that  $v_I \geq v_G$ . 

Although simple, Lemma 3 is actually an important result. If, contrary to Lemma 3,  $v_I = v_G$ , then torture will not change the composition of the pool of denying defendants. That is, given any evidence drawn, the probability that the denying defendant is guilty is exactly the same as the magistrate's prior. In this case, torture will be necessarily inferior to an evidence-based system, as the magistrate can directly convict the defendant, which results in exactly the same type I and type II errors but saves the disutility of pain of torture. The informational advantage brought about by torture can be seen clearly by observing that  $v_I \ge v_G$  implies  $\hat{q}_g > q_g'$  and  $q_i' > \hat{q}_i$ . In other words, compared with the evidencebased system, in the torture system the magistrate is more sure that the suspect is guilty when he draws g and is more sure that the suspect is innocent when he draws i.

Lemma 3 also implies that  $\hat{q}_g > q > \hat{q}_i$ . That is, the magistrate's posterior that the defendant is guilty, after he draws g(i), is greater (smaller) than the prior. Therefore, since the magistrate will not torture under the prior belief (by A2), he will not torture when he draws *i* either. Thus, we have the following result:

**Lemma 4**. If the magistrate draws i, then the suspect is released.

*Proof.* By the first inequality of A2, we know that  $(1-q)L_1 + \frac{kT}{r} > qL_2$ . Since  $q > \hat{q}_i$ , it follows that  $(1 - \hat{q}_i)L_1 + \frac{kT}{r} > \hat{q}_iL_2$ . That is, under the posterior  $\hat{q}_i$ , the expected loss of the magistrate is greater if he tortures the suspect. This implies that the magistrate will release the suspect if he draws i.

By Lemma 4, the magistrate will torture with a positive probability only if he draws g. As a result,  $x_i$  is always 0, and we will use x as a shorthand form of  $x_g$ . That is, x now denotes the probability of torture when the magistrate draws g.

### 4.3 The equilibrium

Given the behavior of the suspect derived in the previous section, the magistrate chooses the value of *x* to minimize his expected loss:

$$\min_{x} W^{T}(x; q, \nu_{G}, \nu_{I}) \equiv x[kT + r(1 - \hat{q}_{g})L_{1} + (1 - r)\hat{q}_{g}L_{2}] + (1 - x)\hat{q}_{g}L_{2} 
= x(kT + r(1 - \hat{q}_{g})L_{1}) + (1 - rx)\hat{q}_{g}L_{2}.$$
[2]

The first term in eq. [2] is the expected loss from torture (the pain of torture plus the cost of a type I error). The second term is the expected loss from a type II error. The equilibrium for the game of torture can then be characterized by  $(v_I^*, v_G^*, x^*)$  in the following proposition:

**Proposition 1.** If the magistrate draws i, he will release the suspect. If he draws g, then:

- If  $\frac{q_G}{q_I} > \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ , then: [1]
  - (i) when  $\frac{P}{rP+T} < q_G$ , then  $v_I^* = 1$ ,  $v_G^* = \underline{v}$ , and  $x^* = \frac{P}{a_G(rP+T)}$ ;
  - (ii) when  $\frac{P}{rP+T}=q_G$ , then  $v_I^*=1$ ,  $v_G^*\in[\underline{v},1)$ , and  $x^*=1$ ;
  - (iii) when  $\frac{P}{rP+T} > q_G$ , then  $v_I^* = v_G^* = 1$ , and  $x^* = 1$ ; where  $\underline{v} = \frac{q_I(1-q)(L_1+kT/r)}{q_Gq(L_2-kT/r)}$ .
- If  $\frac{q_G}{q_I} = \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ , then:
  - (i) when  $\frac{P}{rP+T} < q_G$ , then  $v_I^* = v_G^* = 1$ , and  $x^* \in \left[0, \frac{P}{q_G(rP+T)}\right]$ ;
  - (ii) when  $\frac{P}{rP+T} \ge q_G$ , then  $v_I^* = v_G^* = 1$ , and  $x^* \in [0,1]$ .
- If  $\frac{q_G}{q_I} < \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ , then  $v_I^* = v_G^* = 1$ , and  $x^* = 0$ .

In the case when  $\frac{P}{rP+T} \leq q_I$ , there also exists a pooling equilibrium in which  $v_I^* = v_G^* = 0$ , and  $x^* \in [\frac{P}{a_I(rP+T)}, 1]$ . This is supported by the posterior belief of the magistrate that any non-confessing suspect is guilty with probability  $q^* \in [\frac{L_1 + kT/r}{L_1 + L_2}, 1]$ .

Broadly speaking, there are two kinds of equilibria. The first is a pooling equilibrium in which both types of defendants confess. This is supported by the magistrate's (out-of-equilibrium) belief that a non-confessing suspect is highly likely to be guilty. Under this belief, he tortures with so high a probability that both types of suspect confess before investigation. The second, and more interesting, type of equilibrium is one in which the innocent type never confesses ( $v_I^* = 1$ ). The magistrate's probability of torturing,  $x^*$ , depends on the denying probability of the guilty,  $v_G^*$ . The higher its value, the greater his posterior that a denying suspect is guilty and, therefore, the greater the chance that he will torture. Figure 1 plots the non-pooling equilibrium as a function of the parameters  $\frac{P}{rP+T}$  and  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ 

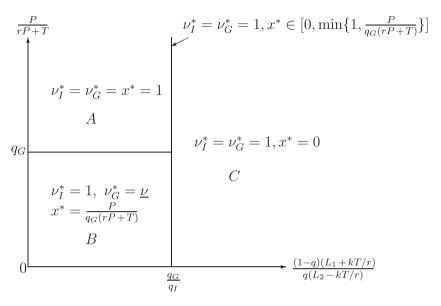


Figure 1: The non-pooling equilibrium.

The value of the vertical axis,  $\frac{P}{rP+T} = \frac{1}{r+\frac{1}{P}}$ , can be viewed as the defendant's relative disutility of penalty versus torture. The greater its value, the less likely it is that the defendant will confess. Indeed, as we move up vertically from any point in Figure 1, the value of  $v_G$  increases. This is, however, also accompanied by an increase of  $x^*$ , as the posterior of a denying suspect being guilty also rises when  $v_G$  increases. When  $v_G$  equals 1, the magistrate's posterior that the suspect is guilty, after he draws g, is so high that he will torture with probability 1. This is the equilibrium outcome in region A of Figure 1. Note that, in this region, all suspects still deny with probability 1 even if they know that they will be tortured with certainty when the magistrate draws g. The reason for this is that P is so large compared to T that they are willing to take the chance of being tortured.

The value on the horizontal axis,  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ , is the expected cost of a type I error relative to that of a type II error. As its value increases, the magistrate will be less inclined to torture. This can be seen by moving from any point in Figure 1 horizontally to the right. In this case, although the value of  $x^*$  remains fixed, the value of  $v_G$  will increase, meaning that the guilty suspect becomes less likely to confess. If the value of  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$  is large enough (specifically, larger than  $\frac{q_G}{q_I}$ ), either because the prior q is low or because the loss of the type I error is large,

**<sup>9</sup>** Equivalently, this means  $\frac{P}{rP+T}$  is greater than  $q_G$ .

the magistrate will be unwilling to torture even if he draws g. In this case,  $x^* = 0$ , and the suspect will never confess. This is the equilibrium outcome in region C of Figure 1. In region B, the value of the legal penalty is so low that the guilty suspect is inclined to deny the guilt. This is, however, tempered by the high probability that the magistrate will torture (since the value of  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$  is low). In the equilibrium, he confesses with a probability greater than 0 but less than 1.

Proposition 1 also implies that the magistrate uses a more lenient rule to convict a defendant under the torture system. Recall that under the evidencebased system, the magistrate convicts the suspect if and only if he draws g. Here, it is still the case that the suspect will be released when i is drawn. However, the magistrate will not necessarily torture (and convict) the suspect even if g is drawn.

Note that of all equilibria, the guilty and the innocent act differently only in region B of the non-pooling equilibrium. That is, only in region B does the torture system force the guilty suspect to confess more readily. This indicates that the torture system can incur a lower social loss than an evidence-based system only in region B, a fact that will be proved in Section 5.

#### 5 Evidence vs torture

The central question that the paper intends to answer is which of the two systems incurs a lower expected cost. Torture, by forcing a suspect to confess regardless of whether he is guilty or not, decreases the chance of the type II error at the expense of both the pain imposed on the suspect and an increase in the type I error. As a result, the relative merit of the two systems critically depends on two factors. The first is the relative size of the losses brought about by type I and type II errors. The second is the accuracy of the magistrate's investigation. In this section, we will first derive a precise condition under which system it incurs a lower loss than the other. Based on this condition, we then show how their relative merit will change in response to an improvement in information revealed in the investigation. Our result is then used to explain certain aspects of the historical evolution of the judicial system.

### 5.1 Welfare comparison

Under the evidence-based system, with probability  $q(1-q_G) + (1-q)(1-q_I)$ the magistrate will draw i, in which case the defendant will be released. Since there will be no type I error, the expected loss is  $q(1-q_G)L_2$ . If he draws g, which occurs with probability  $qq_G + (1-q)q_I$ , the magistrate will convict the defendant. This results in an expected loss of  $(1-q)q_IL_1$ . The total expected loss under the evidence-based system is thus

$$W^{E} \equiv q(1 - q_{G})L_{2} + (1 - q)q_{I}L_{1}.$$
 [3]

The expected loss of the torture system depends on which equilibrium outcome we consider. For the pooling equilibrium, in which  $v_I^* = v_G^* = 0$  and  $x^* \in \left[\frac{P}{q_I(rP+T)},1\right]$ , the expected loss is  $W^T = (1-q)L_1$ . As a result, the difference of the constant of the pooling equilibrium. ence in the expected loss between the two systems is

$$\Delta W \equiv W^T - W^E = (1 - q)L_1 - [(1 - q)q_IL_1 + q(1 - q_G)L_2]$$
  
=  $(1 - q)(1 - q_I)L_1 - q(1 - q_G)L_2.$  [4]

As a result,  $\Delta W \leq 0$  if and only if

$$\frac{1 - q_G}{1 - q_I} \ge \frac{(1 - q)L_1}{qL_2}.$$
 [5]

Inequality [5], however, violates assumption A1. This means that if an investigation is of certain informational value in the sense of A1, the evidencebased system always results in a lower expected loss than the pooling equilibrium in the torture system. The reason for this is as follows. In the pooling equilibrium, the defendant always confesses, and the magistrate never needs to investigate. The torture system thus has not utilized the information that would have been revealed had there been an investigation. Given that the investigation is of informational value, the evidence-based system, as one in which the magistrate's decision is based on evidence, is naturally superior.

For the non-pooling equilibrium, the comparison becomes much more complicated, because the equilibrium outcome depends on parametric configurations. In general, the expected loss under the torture system is

$$W^{T} \equiv (1 - q)[(1 - v_{I}^{*}) + v_{I}^{*}q_{I}x^{*}r]L_{1} + qv_{G}^{*}[(1 - q_{G}) + q_{G}(1 - x^{*} + x^{*}(1 - r))]L_{2}$$

$$+ [v_{I}^{*}(1 - q)q_{I} + v_{G}^{*}qq_{G}]x^{*}kT$$

$$= (1 - q)q_{I}x^{*}rL_{1} + qv_{G}^{*}(1 - q_{G}x^{*}r)L_{2} + [(1 - q)q_{I} + v_{G}^{*}qq_{G}]x^{*}kT,$$
[6]

<sup>10</sup> Note that, under the pooling equilibrium, both types confess in the equilibrium, so that the magistrate does not need to investigate. We thus do not need to distinguish the cases when he draws g and i.

where the equality comes from the fact that  $v_I^*$  is always 1 in the non-pooling equilibrium.

In region A of Figure 1 where  $v_I^* = v_G^* = x^* = 1$ ,  $W^T$  reduces to  $(1-q)q_IrL_1 + q(1-q_Gr)L_2 + [(1-q)q_I + qq_G]kT$ . We then have

$$\Delta W = (1-r)[qq_GL_2 - (1-q)q_IL] + [(1-q)q_I + qq_G]kT > 0.$$
 [7]

The first term is positive because of assumption A1. Therefore, if the suspect always denies and the magistrate always tortures when he draws g under the torture system, the evidence-based system is superior. The reason for this is as follows. Under the equilibrium in region A, the magistrate will torture a suspect if and only if he draws g. With probability r, the suspect confesses after being tortured, so that the outcomes are exactly the same in these two systems. However, the torture system is inferior by the magnitude of the expected cost of torture. With probability 1-r, the suspect still denies after being tortured, and is released so that torture is a waste. As long as the investigation is informative, the evidence-based system is still superior.

In region *C* of Figure 1 where  $v_I^* = v_G^* = 1$  and  $x^* = 0$ ,  $W^T$  reduces to  $qL_2$ . Thus,

$$\Delta W = qL_2 - (1-q)q_IL_1 - q(1-q_G)L_2 = qq_GL_2 - (1-q)q_IL_1.$$
 [8]

Therefore,  $\Delta W > 0$  if and only if

$$\frac{(1-q)L_1}{qL_2} < \frac{q_G}{q_I},\tag{9}$$

which always holds by A1. This implies that if the suspect always denies and the magistrate never tortures under the torture system, the evidence-based system incurs a lower social loss. The reason for this is actually very simple. In the case when  $v_I^* = v_G^* = 1$  and  $x^* = 0$ , a system based on torture actually does nothing. Suspects are routinely summoned to the court and released. Given that the investigation is of informational value (i.e. A1), a system that convicts a suspect based on evidence is naturally superior.

All the three cases above share a common feature: Information that is available through investigation is not utilized in the torture system in an efficient way. Either all types of suspect confess, so that investigation is not needed, or all who deny are released, or the magistrate convicts the suspect in exactly the same fashion as the evidence-based system. Since information offered by investigation is valuable, and since an evidence-based system does utilize this piece of information, the torture system always incurs a greater loss. To summarize, if the magistrate's decision is not altered by the outcomes of the investigation in a torture system, or the criterion to convict a suspect is the same as in the evidence-based system, a torture system is inferior.

The more interesting case occurs when the equilibrium outcome falls in region B of Figure 1, i.e. when the evidence collected will affect the magistrate's decision in a way different from the evidence-based system. In this case, the difference in the expected loss is

$$\begin{split} \Delta W &= -(1-q)(1-x^*r)q_IL_1 + q[\underline{\nu}(1-q_Gx^*r) - (1-q_G)]L_2 + [(1-q)q_I + \underline{\nu}qq_G]x^*kT \\ &= (1-q)q_Ix^*r(L_1+kT/r) - qq_Gx^*r\underline{\nu}(L_2-kT/r) - (1-q)q_IL_1 - q(1-q_G-\underline{\nu})L_2 \\ &= -[(1-q)q_IL_1 + q(1-q_G-\underline{\nu})L_2], \end{split}$$

where the last equality in eq. [10] comes from the fact that the first two terms after the second equality sum to 0. Note that eq. [10] is an increasing function of  $\underline{v}$ . Moreover,  $\Delta W < 0$  when  $\underline{v} = 0$  and  $\Delta W > 0$  when  $\underline{v} = 1$ . Therefore, there exists a  $v^0 \equiv \frac{(1-q)q_1L_1}{qL_2} + 1 - q_G \in (0,1)$  such that  $\Delta W > 0$  if and only if  $\underline{v} > v^0$ . However, since  $\underline{v} = \frac{q_L}{q_G} \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ , we know that  $\Delta W > 0$  if and only if

$$\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)} \ge v^0 \frac{q_G}{q_I}.$$
 [11]

We can now summarize the results of our comparison in Figure 2. The curve  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}=\frac{q_G}{q_I}v^0$  partitions region B in Figure 1 into two parts,  $B_1$  and  $B_2$ . Configurations lying in  $B_1$  ( $B_2$ ) result in equilibria having a lower (higher)

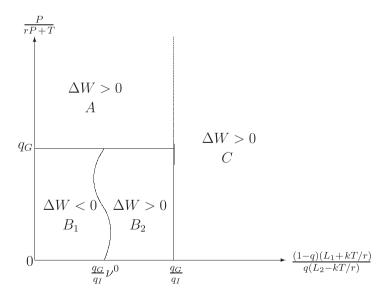


Figure 2: Expected loss comparison between two systems.

expected loss in the torture system than the evidence-based system. Therefore, among all possible equilibria, the torture system is superior only when the configurations of the parameters lie in region  $B_1$ , in which the values of both  $rac{P}{rP+T}$  and  $rac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$  are lower. In other words, a torture system is superior only when the relative disutility of legal penalty to torture is low, and the expected cost of a type I error relative to that of a type II error is low. Recall that torture imposes two kinds of cost on the society. One is the pain of torture and the other is the increase in the type I error. Thus, only when the type II error imposes a serious cost in the magistrate's eyes, and the legal penalty is not great enough to be deterring, can the magistrate be justified in resorting to torture. However, if one of the conditions is not met, then torture is more costly, and a system based solely on evidence will be better.

#### 5.2 Improvement in informativeness of investigations

In this section, we investigate how the behavior of the suspect and the magistrate will change in response to an improvement in the information revealed during the investigation. As can be seen in Proposition 1, an important parameter that influences the behavior of both the magistrate and the suspect is  $q_G/q_I$ . Since  $q_G(resp. 1 - q_I)$  is the probability that the magistrate will draw g(resp. i) when the defendant is guilty (resp. innocent), an increase in  $q_G$  or a decrease in  $q_I$ , (and therefore an increase in  $q_G/q_I$ ) implies that more precise information is released during the investigation. The value of  $q_G/q_I$  is therefore a measure of the informativeness of the investigation. We will say that the investigation becomes more informative if  $q_G$  increases or  $q_I$  decreases, with at least one strictly so.

We will show an interesting result of a non-monotonic relationship between the informativeness of investigation (in terms of  $q_G/q_I$ ) and the attractiveness of the torture system (the occurrence of  $\Delta W < 0$ ): when little information is revealed during investigation, the torture system performs even better when information is improved; however, as investigation becomes sufficiently informative (i.e.  $q_G/q_I$  has passed some threshold), then further improvement in the information will begin to lend an advantage to the evidence-based system. Eventually, the torture system is dominated.

The reasoning of this result is the following. When the investigation becomes more informative, the outcomes in both systems will be affected. In the evidence-based system, both inequalities in assumption A1 become more likely to hold. Consequently, the magistrate is more likely to release a suspect when he draws *i*, and more likely to convict him when he draws *g*. Moreover, the expected losses from a type I error (when he convicts a suspect) and a type II error (when he releases a suspect) both decrease. This can be easily seen from eq. [3]: As the investigation becomes more informative,  $W^E$  will decrease. In particular, when  $q_G$  approaches 1 and  $q_I$  approaches 0, the social loss of the evidence-based system approaches 0.

In the torture system, both  $\hat{q}_g$  and  $\hat{q}_i$  will increase as the investigation becomes more informative. Furthermore, the line  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}=\frac{q_G}{q_I}$  will shift to the right. In this case, region C will retract, and regions A and B will expand. Therefore, the magistrate will be more likely to torture on the one hand, and the guilty suspect will be more likely to confess on the other. Note that as  $q_G$ approaches 1 and  $q_I$  approaches 0,  $\underline{y}$  will approach 0 and the probability that the magistrate will torture,  $x^*$ , will approach  $\frac{P}{rP+T}$ . This resembles a separating equilibrium. However, unlike the evidence-based system, what separates the innocent and the guilty is not the evidence, but the fact that the innocent always denies and the guilty always confesses. This relies on the magistrate's commitment that his probability to torture is bounded away from 0.

It can be shown that  $\frac{q_G}{q_I}v^0$  is a decreasing function of  $q_I$ , meaning that as  $q_I$ becomes smaller, the dividing curve between  $B_1$  and  $B_2$  will shift to the right. In other words, the torture system will gain its advantage as  $q_I$  decreases. Moreover,  $\frac{q_G}{q_I}v^0$  is increasing (decreasing) in  $q_G$  if

$$q_G < (>) \frac{q_I \frac{1-q}{q} \frac{L_1}{L_2} + 1}{2}.$$
 [12]

That is, as the value of  $q_G$  increases from a small value,  $\frac{q_G}{q_I}v^0$  will first increase, and then decrease as the value of  $q_G$  passes the threshold value on the righthand side of eq. [12]. Consequently, an increase in  $q_G$  will favor the torture system when  $q_G$  is small, and favor the evidence-based system when it is large. Note that the smaller the value of  $q_I$ , the more likely it is that  $q_G$  is greater than the right-hand side of eq. [12]. Therefore, at the same time the magistrate becomes more competent in identifying the innocent suspect, an increase in  $q_G$  is more likely to lend advantage to the evidence-based system. In particular, when  $q_G$  is sufficiently large, then  $\Delta W \geq 0$  regardless of the value of  $q_I$ . This can be seen from the following fact: if  $q_G$  is sufficiently close to 1, then  $v^0$  is close to  $\frac{(1-q)q_lL_1}{qL_2}$ , so that it is always the case that  $\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)} \geq v^0 \frac{q_G}{q_l}$ . In other words, as the magistrate's ability to identify a guilty suspect through investigation passes a certain threshold, there will be no more room for torture: a system based on evidence is always less costly. When this is true, the evidence-based system will incur a lower social cost in all possible equilibria, and will be unambiguously better and eventually guarantee its domination.

In Section 7, we review the historical development of judicial torture, which appears to be consistent with our theoretical results. First note that although medical expertise had started to be used in judicial cases since the thirteenth century (so that the precision of investigation had improved), there was actually an increase in the use of torture. This is consistent with our prediction that an improvement in the technology of investigation lends advantage to torture, when the investigation is relatively uninformative. Only after the sixteenth century (the beginning of the use of modern forensic science) did the practice of torture start to decline. This is also consistent with our prediction that as the technology of investigation becomes sufficiently advanced, an increase in informativeness will give an advantage to an evidence-based system. Finally, at the end of nineteenth century, when more advanced testing methods (like fingerprints and blood type identification) started to be accepted in court and when the systematic use of scientific investigation for legal use became common, a system based on evidence dominated thereafter.

#### 5.3 An illustrative example

In this section, we use a simple example to demonstrate the non-monotonicity result found in the previous section. Consider the symmetric case where q = 1/2and  $L_1 = L_2 = L$ . That is, it is believed that the suspect is guilty or innocent with equal probability a priori, and the costs of type I and type II errors are equally important to the magistrate.

We have shown that the torture system is worse than the evidence-based system in the pooling equilibrium and in regions A and C of the non-pooling equilibrium. We only need to discuss region *B* of the non-pooling equilibrium where the torture system can possibly win. Based on eq. [10], we have

$$\Delta W = -\frac{L}{2}(q_I + 1 - q_G - \underline{v}) = -\frac{L}{2}\left(q_I + 1 - q_G - \frac{q_I}{q_G} \cdot \frac{L + kT/r}{L - kT/r}\right).$$

Let  $\ell = \frac{L + kT/r}{L - kT/r}$ . Therefore,  $\Delta W < 0$  if and only if

$$q_G + \frac{q_G}{q_I} - \frac{(q_G)^2}{q_I} \equiv b(q_I, q_G) > \ell.$$
 [13]

 $\ell$  represents the relative cost between torturing and not torturing the suspect. Obviously,  $\ell > 1$ . On the other hand,  $b(q_I, q_G)$  represents the relative benefit of information elicitation by confession via torturing the suspect. It is easy to see that: (1)  $b(q_I, q_G) > 1$  since  $q_I < q_G < 1$ , and  $b(q_I, q_G)$  is concave in  $q_G$ ; (2)  $b(q_I,q_G) \rightarrow 1$  when  $q_G \rightarrow q_I$  and  $q_G \rightarrow 1$  and (3)  $b(q_I,q_G)$  reaches its peak

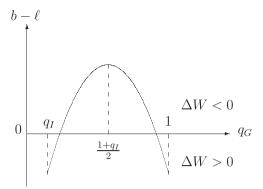


Figure 3: The net benefit of torture.

when  $q_G = \frac{1+q_I}{2}$ . The relationship between the net benefit of torture,  $b - \ell$ , and the level of  $q_G$  (given  $q_I$ ) is drawn in Figure 3.

As shown in Figure 3, when  $q_G$  approaches  $q_I$ ,  $b - \ell < 0$ , which means that the torture system loses to the evidence-based system. This is because investigation does not reveal too much information, so that it is very likely that the guilty suspect will deny the crime (i.e. v is close to 1). In this case, torture does not bring much advantage but causes a social cost. On the other hand, when  $q_G$ approaches 1, again,  $b - \ell < 0$ , because investigation has already done a good job in separating the guilty from the innocent, so that torture cannot outperform the evidence-based system. It is when  $q_G$  is at some intermediate level that the torture system can become better than the evidence-based system. The relationship between  $b-\ell$  and  $q_G$  is then hump shaped: when  $q_G$  is small, an increase in  $q_G$  will even improve the benefit from using torture; however, once it passes the threshold  $\frac{1+q_I}{2}$ , torture becomes less attractive as  $q_G$  becomes higher, and eventually becomes inferior to the evidence-based system.

# 6 Torturing plaintiffs or witnesses

Judicial torture in pre-modern societies was legitimately applied almost universally to suspects, witnesses and plaintiffs for various reasons before the midnineteenth century. The accuser or the plaintiff would undergo an ordeal to substantiate his veracity or be tortured when he was unable to make good his accusation. In Roman law, the accuser could be exposed to the lex talionis (law of retaliation) in case he failed to prove the justice of the charge (Lea 1971, 333). In primitive Russian laws, the accuser was obliged to undergo the ordeal of a red-hot iron if he could not substantiate his case with witnesses. Archbishop Hincmar of Rheims of the ninth century required that cases of complaints against priests be supported by seven witnesses, of whom one must be tortured to prove the truth of his companions' oath, as a wholesome check upon perjury and subornation (Lea 1971, 290).

The same thing could be observed in China. The legal code of the Qing dynasty clearly allowed torture to be used not only on the accused, but also on "secondary suspects and to witnesses as well as to principals" (Bodde and Morris 1967, 97). In fact, the tradition could be traced back to the Tang code of AD 653 that constituted the basis for all subsequent imperial law codes. According to Tang law, if the accused insisted on his innocence even after having suffered the maximum amount of torture allowed by the law, the plaintiff would in turn be tortured (Shen 1985, 511). Moreover, since at least the Tang dynasty, a plaintiff who bypassed the immediate authority and presented his complaints to higher administrative levels was required to undergo torture before the examination of the case (Xue 1998, 636-9), again as a check upon perjury and subornation, or as a measure to discourage such actions. During the early years of the Ming dynasty (1368-1644), frequent complaints brought directly to the capital prompted the government to implement an extremely severe punishment - banishment to the frontier - on plaintiffs, even at times on those who could substantiate their cases, to control the number of such acts (Shen 1985, 1142).

In this section, we analyze the puzzling phenomenon that the plaintiffs or the witnesses are tortured in the judicial procedure. Since the basic logic behind our argument is the same for torturing the plaintiffs and the witnesses, we will consider only the former. Our basic model in Section 2 is modified in the following way. Suppose that, before the magistrate accepts a case, he can torture the plaintiff, ostensibly to "verify" or "test" whether the plaintiff is telling the truth. Once tortured, the disutility for the plaintiff is  $T_p$ , and it also yields a loss to the magistrate,  $k_pT_p$ . The events then follow as described in the model mentioned before.

In the previous sections, we have assumed that the prior of the magistrate that the suspect is guilty is q. Here, in order to discuss how torture changes the plaintiff's incentives to come to the court, and thus the magistrate's prior, we assume that q is determined by the cases that plaintiffs bring into the court. Specifically, suppose that every plaintiff has a belief regarding how likely the defendant is to be guilty, represented by a probability  $\rho$ . The distribution of plaintiffs' beliefs is given by the density function  $f(\rho)$ . Hence, the prior of the magistrate that a defendant is guilty, if all plaintiffs go to the court, will be  $\int_0^1 \rho f(\rho) d\rho$ . Thus, what we call the prior of the magistrate in the previous sections is actually the mean value of the beliefs of the plaintiffs who actually enter the court.

Assume that a plaintiff obtains a utility Y > 0 when a guilty defendant is convicted. If the guilty suspect is wrongfully released, or the innocent defendant is wrongfully convicted, the plaintiff is assumed to obtain a negative utility -y, v > 0. The plaintiff's utility is 0 if an innocent suspect is released. Therefore, given the probability that the magistrate will torture the plaintiff,  $\varphi$ , the expected utility for the plaintiff with a belief  $\rho$ , when he comes to the court, is

$$\rho[(1-v_G^*)+v_G^*q_Gx^*r]Y-\rho v_G^*(1-q_Gx^*r)y-(1-\rho)[(1-v_I^*)+v_I^*q_Ix^*r]y-\varphi T_p.$$
[14]

The first term in eq. [14] is the expected utility that a guilty suspect is convicted. The second and third terms are the expected losses from type II and type I errors, respectively. The fourth term is the expected cost of torture. If, on the other hand, the plaintiff decides not to come into the court, his expected payoff is simply  $-\rho y$ .

To see how torturing the plaintiff affects the equilibrium of the torture game, we again need to discuss the two types of equilibrium in Proposition 1 separately, i.e. the pooling and the non-pooling equilibria. The details of the analysis are relegated to Appendix B. The key driving force is that, by torturing the plaintiff, the incentive of the plaintiff to come to the court is reduced. Those plaintiffs who are less sure of the defendants' guilt are deterred. Consequently, the average quality of cases brought to the court (which is represented by  $q(\varphi)=\int_{
ho(\varphi)}^{1}rac{
ho f(
ho)}{1-F(
ho(\varphi))}d
ho)$  is increased in the sense that the suspect has a higher probability of being guilty.

The main intuition for the result that torturing the plaintiff or witness can sometimes be optimal is the following. In the pooling equilibrium, both types of suspect confess. Thus, the expected loss under torture is  $W^T(\varphi) = (1 - q(\varphi))$  $L_1 + \varphi k_p T_p$ . If the magistrate tortures the plaintiff by increasing the value of  $\varphi$ ,  $q(\varphi)$  will be increased so that the expected cost related to the type I error will be reduced. Therefore, when  $L_1$  is large enough, the benefit of raising the quality of cases entering into court will outweigh its cost  $k_pT_p$ , so that torturing the plaintiff can be optimal.

In the non-pooling equilibrium, the expected loss is  $W^T(\varphi) = (1 - e^{-1})^T (\varphi)$  $q(\varphi) q_I x^* r L_1 + q(\varphi) v_G^* (1 - x^* q_G r) L_2 + [(1 - q(\varphi)) q_I + q(\varphi) v_G^* q_G] x^* k T + \varphi k_p T_p$ . If the magistrate never tortures the suspect (i.e.  $x^* = 0$ , or region C in Figure 1), the expected loss will be  $qL_2 + \varphi k_p T_p$ . Obviously, it is optimal to choose  $\varphi^* = 0$ . This implies that if the magistrate decides to torture the plaintiff or the witness, he will torture the suspect as well. On the other hand, if the magistrate decides to torture the suspect with a positive probability (i.e.  $x^* > 0$ , or regions A and B in Figure 1), by torturing the plaintiff which further raises q, the cost of the type II error may even be reduced since it would be more likely for the guilty suspect to confess (i.e.  $\frac{\partial v_c^*}{\partial a} \leq 0$ ). The magistrate thus trades off the loss resulting from torturing the plaintiff with the benefit from reducing both the type I and type II errors. If the benefit surpasses the loss, the magistrate will torture the plaintiff.

We also find that when either *Y* or *y* is larger, which means that the plaintiff has a stronger sense of "justice" in that he derives (suffers) a greater utility (disutility) in having a guilty (innocent) suspect convicted, then torture will be less likely to deter him from coming into court. Thus, the screening function of torture is smaller, so that it will be less likely to be used.

More importantly, when the investigation is less informative (i.e.  $q_G/q_I$  is smaller), the screening effect of torture will be stronger. In this case, torturing the plaintiff will occur more frequently. This means that there is a negative relationship between the informativeness of the investigation and the tendency to torture the plaintiff. Moreover, since the magistrate is also more likely to torture the suspect when  $q_G/q_I$  is relatively low, the two practices (torturing the suspect and torturing the plaintiff or the witness) generally go hand in hand.

#### 7 The historical evolution of torture

In this section, we would like to review the historical evolution of torture and show that it is consistent with our theoretical predictions.

European countries do not have exactly the same history of judicial torture, as each has a different judicial tradition and religious culture that greatly affects legal application. However, from the specific cases of England and France, one may construct a picture reflecting the general history of judicial torture in Europe: the application of judicial torture began to soar from around the thirteenth century11 onward and peaked in the sixteenth century. It declined thereafter and was nominally abolished in most European countries toward the end of the eighteenth and early nineteenth centuries. In England, "references to its use (use of judicial torture) in the earlier years are remarkably scanty, but during the fifteenth, sixteenth and seventeenth centuries, the evidence is abundant...(J)udicial torture reached its greatest ecumenicity in the reign of Elizabeth (1533–1603)" (Scott 2003, 88–9). In France, in a southern town such as Toulouse where the tradition of Roman law was strong, judicial torture "was first

<sup>11 &</sup>quot;Torture was introduced for the express purpose of extracting confession," being authorized by Pope Innocent in a Bull issued in 1252 (Scott 2003, 66).

employed... during the 13th century by the town consuls," and it "reflected a new methodology and a new epistemology" (Silverman 2001, 5-7). The decline in the actual use of judicial torture in France in general seemed to begin from the mid-sixteenth century to the seventeenth century, depending on the region, and lasted until its nominal abolition in 1788 (Silverman 2001, 18). The pattern corresponds roughly to the one observed in England. By 1850, the movement for the abolition of torture swept most parts of Europe. Furthermore, "by the closing decades of the 19th century it was widely thought that torture was a barbaric practice that belonged to history" (Evans and Morgan 1998, 12–13). Although recent scholarship still casts doubt on the real disappearance of judicial torture in Europe by the end of the nineteenth century, as evidence suggests the persistence of the illegal use of judicial torture in a democratic country like the United States even in the 1920s (Evans and Morgan 1998, 14), such evidence shows that torture was no longer the norm. The decline of judicial torture was therefore a continuous process beginning from the sixteenth century up to the early twentieth century. It was legally and ideologically condemned by the end of the eighteenth century, and, in practice, it may have persisted up to the turn of the twentieth century.

The decline of judicial torture in modern Europe has been the subject of a number of scholarly works in the past few decades. The conventional historical account of the abolition movement placed great weight on the philosophic goals of the Enlightenment, which is considered to be the driving force behind the abolition of torture. <sup>12</sup> Langbein proposes a more specific albeit controversial explanation based on changes within the penological history itself. He suggests that judicial torture became increasingly useless when a new law of proof emerged around the sixteenth century that no longer required a strict standard of proof, accompanied by less rigorous punishments, a development he calls the "evidentiary revolution" (Langbein 1977, 1983). Langbein's thesis was criticized by another law specialist M. Damaška, especially on the significance of the "evidentiary revolution" in the sixteenth century. For Damaška, judicial torture still had a raison d'être in Europe during and after the sixteenth century (Damaška 1978).

More recently, historian Lisa Silverman has proposed another explanation for the final abolition of judicial torture in eighteenth-century France, partially rehabilitating the humanist theory rejected by Langbein. For Silverman, there was a "dramatic paradigm shift" in the way abolitionists of the Enlightenment understood the relation between judicial torture (pain), truth and the body. Bodily pain was no longer believed to be able to produce truth (Silverman

<sup>12</sup> See discussion in Silverman (2001), and the references therein.

2001). One reason Silverman does not find satisfaction in Langbein's theory is that long after the "evidentiary revolution," judicial torture was still practiced in many parts of Europe, a point also raised in Damaška's review.

This historical evolution of judicial torture in Europe corresponds meaningfully with our model if we look at another factor that might explain in a more satisfactory way the rise and decline of judicial torture: the employment of medical science in law. There are indications of the application of medical expertise in judicial cases in northern Europe from the tenth century onward, and it was explicitly mentioned in Norman law from the early thirteenth century onward. Between the thirteenth and the sixteenth centuries, the trend exhibited a stable increase until the publication in 1562 of the first judicial postmortem in France by Ambroise Paré, although the conventional wisdom is that the first comprehensive work on forensic medicine, De relationibus medicorum libri quatuor, was published by Fortunato Fedele (1550–1630) in 1602. An increasing number of works of the kind were published throughout the seventeenth and eighteenth centuries with the first serial devoted to forensic medicine published in 1755. These early European publications on forensic medicine indicate that the practice of judicial forensic medicine was already common throughout the sixteenth century, if not earlier. Later developments in the science of fingerprinting (from the late seventeenth century), physical matching (late eighteenth century), anthropometry (late nineteenth century) and so on were the intensification of the same trend. More significantly, the systematic application of modern scientific technology in criminal investigations began in the last decades of the nineteenth century. Most notable was the new classification system created by E. Henry in 1896 enabling fingerprints to be easily filed, searched and traced. It was quickly used worldwide and is still used today.<sup>13</sup> Henry himself was appointed assistant administrator of Scotland Yard in 1901 in charge of the Criminal Investigation Department, and forced the adoption of fingerprint identification.<sup>14</sup> New York State in the United States followed in 1903.<sup>15</sup> Most European countries established crime detection laboratories as government or university units in the first quarter of the twentieth century (Parker 1983, 432). In other words, there were two critical points in the development of forensic medicine in criminal investigation: first, the sixteenth century marked the beginning of modern forensic medicine for legal uses (and the "evidentiary revolution"). Second, the end of the nineteenth century marked a paradigm shift in such technology and its systematic application in criminal investigation (Nemec

<sup>13</sup> See "The History of Fingerprints," http://onin.com/fp/fphistory.html

<sup>14</sup> Wikipedia, http://en.wikipedia.org/wiki/Edward\_ Henry.

<sup>15</sup> See "Forensic Science Timeline," http://www.forensicscience.org

1968, 5–15; Saferstein 2001, 3; Parker 1983, 430). At the same time, the sixteenth century represents the peak of the legal employment of torture in most parts of Europe. The decline, abolition and the ultimate disappearance of judicial torture in democratic countries would take more than 300 years.

### 8 Conclusion

This paper applies the economic theory of information to analyze the nature and consequences of judicial torture. The theory also suggests an explanation for the rise and fall of torture in Europe. It is shown that if the magistrate aims to balance type I and type II errors in judgment, and if during the investigation little information is revealed, then a system based on torture functions better than one based on evidence. This mainly comes from the ability of torture to force the guilty defendant to confess with a higher probability than the innocent. It is also shown that the information advantage of a torture system will be more pronounced as the precision of investigation rises from a low level. However, as its precision passes a certain threshold, the advantage of torture decreases relative to a system based on evidence. This may explain an increase in the practice of torture before the sixteenth century on the one hand, and a transition from a torture system to an evidence-based system starting from the mid-eighteenth century in Europe on the other. Furthermore, torturing the plaintiffs might also be an equilibrium outcome, as it helps to screen the cases so that only those with greater merits enter the court. Again, this is more likely to occur when the quality of information revealed during the investigation is low.

There have been two major theories to explain the abolition of torture, one based on humanist concern and the other based on legal change. In this paper we propose an alternative theory based on the progress of technology. We hasten to add that our theory does not mean to replace, or even compete with, the existing explanations. All three factors might contribute to the decline of torture, and in this sense they are complementary to each other. In fact, the humanist theory and the legal theory can both comfortably fit into our framework. The reason why there had been a change in the rule of proof might be precisely because the improvement in the technology of evidence collection has been such that a system based solely on evidence was found to reach more precise verdicts. On the other hand, a change in the attitude of the general public toward torture is equivalent to an increase of  $L_1$  in our model, which shifts the curve dividing  $B_1$  and  $B_2$  to the left, and therefore increases the cost of torture. From this viewpoint, the contribution of this paper might be seen as

providing a unifying framework in which the humanist, legal, and technology theories are all important components in explaining the rise and fall of judicial torture.

The paper's results also have some political implication. Western countries condemn regimes that use torture. To reduce the prevalence of torture in those less-developed countries, the developed countries can provide forensic technology in the form of "judicial development aid," instead of intervention or sanctions. However, given the non-monotonic result of torture's effectiveness, we suggest that it is important for the technology to be advanced enough to avoid its adverse effect.16

# Appendix A: Proof of Proposition 1

First note that an innocent suspect will confess before being tortured with a positive probability only if  $P < q_I(rP+T)x$ , i.e. the loss of confessing is smaller than that of denying. The similar condition for a guilty suspect is  $P \leq q_G(rP+T)x$ . Also, given the posterior  $\hat{q}_g$ , the magistrate has a positive probability to torture only if  $\hat{q}_g L_2 \geq (1 - \hat{q}_g) L_1 + \frac{kT}{r}$ , or, equivalently,  $\frac{q_G}{q_I} \ge \frac{(1-q)v_I(L_1+kT/r)}{qv_G(L_2-kT/r)}$ 

We first focus on the equilibrium in which  $v_I^* > 0$ .

[1] Suppose 
$$\frac{q_G}{q_I} > \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$$
:

First of all, we only have to consider the case x > 0 in equilibrium. The reason is as follows. If x = 0, then  $v_I = v_G = 1$ , because it is always the case where  $P > q_I(rP + T)x$  and  $P > q_G(rP + T)x$ . However, if this were the case, then x = 1due to the assumption  $\frac{q_G}{q_I}>\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$ , which is a contradiction. Secondly, if in equilibrium,  $x \in (0,1)$ , then it must be the case that  $v_I = 1$  and  $v_G \in (0,1)$ . To see this, recall that we have already shown that  $v_I \geq v_G$ . If  $v_I = v_G = 1$ , then again it must be the case that x = 1, which violates the assumption  $x \in (0,1)$ . If  $v_I = 1$ , or  $v_I \in (0,1)$  and  $v_G = 0$ , then x = 0, which is also a contradiction.

There are three further cases to consider: (1) If  $\frac{P}{rP+T} < q_G$ , it is not possible that x = 1 in equilibrium. The reason is the following. If x = 1, since  $P < q_G(rP + T)x$ ,  $v_G = 0$ . It follows that x = 0 as long as  $v_I > 0$ , which is a

<sup>16</sup> We are grateful to Burkhard Schipper for suggesting this political implication for the model.

contradiction. Thus,  $x \in (0,1)$ . Since the magistrate is using a mixed strategy, it must be the case that he is indifferent between torturing and releasing the suspect. That is, it must be the case that  $\frac{q_G}{q_I} = \frac{(1-q)v_I(L_1+kT/r)}{qv_G(L_2-kT/r)}$ . As we have already shown,  $x \in (0,1)$  also implies  $v_G \in (0,1)$ . Again, the guilty suspect is indifferent between confessing and denying, so  $P = q_G(rP+T)x$ . Thus, in equilibrium,  $v_I^* = 1$ ,  $v_G^* = \frac{q_I(1-q)(L_1+kT/r)}{q_Gq(L_2-kT/r)} \equiv \underline{v}$  and  $x^* = \frac{P}{q_G(rP+T)}$ . (2) If  $q_G = \frac{P}{rP+T} > q_I$ , then  $P = q_G(rP+T)$  and  $P > q_I(rP+T)$ . The inequality above implies that  $v_I = 1$ . If x = 1, then  $v_G \in (0,1)$ . Since by assumption  $\frac{q_G}{q_I} > \frac{(1-q)(L_1+kT/r)}{q_{VG}(L_2-kT/r)}$ , it must be that  $v_G > \underline{v}$ . If  $x \in (0,1)$ , then  $v_G = 1$ , which will result in x = 1, a contradiction. Therefore, the equilibrium is  $v_I^* = 1$ ,  $v_G^* \in (\underline{v},1)$ , and  $x^* = 1$ . (3) If  $\frac{P}{rP+T} > q_G$ , then both types of suspect prefer to deny for any x, so  $v_I = v_G = 1$ . It follows that  $x^* = 1$ .

[2] Suppose 
$$\frac{q_G}{q_I} = \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$$
:

First, if x=0, then it must be that  $v_G=v_I=1$ . The fact that  $\frac{q_G}{q_I}=\frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$  thus implies that the magistrate is indifferent between torturing and releasing the suspect. That is,  $x^*=0$  and  $v_I^*=v_G^*=1$  constitute an equilibrium. Next, if x>0, then it must be the case that to torture is at least as good as to release for the magistrate, which means  $\frac{q_G}{q_I}\geq\frac{(1-q)v_I(L_1+kT/r)}{qv_G(L_2-kT/r)}$ . Since  $v_I\geq v_G$ , this is possible only if  $v_I=v_G$ . However, if  $v_I=v_G<1$ , then it must be that  $P\leq q_I(rP+T)x$  and  $P< q_G(rP+T)x$  (note that  $q_G>q_I$ ). This implies that  $v_G=0$  and thus,  $v_I=0$ . This violates the case that we set out to consider  $(v_I^*>0)$ . This means that if x>0,  $v_G=v_I=1$ . There are thus two cases to consider. (1) Let  $\frac{P}{rP+T}\geq q_G$ . Then  $P\geq q_G(rP+T)x$  and  $P>q_I(rP+T)x$  for any  $x\in[0,1]$ . Consequently,  $x^*\in[0,1]$  and  $v_I^*=v_G^*=1$  constitutes an equilibrium. (2) Let  $\frac{P}{rP+T}<q_G$ , then in order for  $v_I^*=v_G^*=1$  to hold, it must be the case that  $P\geq q_G(rP+T)x$ , i.e.  $x\in[0,\frac{P}{q_G(rP+T)}]$ . In other words,  $v_I^*=v_G^*=1$  and  $x^*\in[0,\frac{P}{q_G(rP+T)}]$  constitutes an equilibrium.

[3] Suppose 
$$\frac{q_G}{q_I} < \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$$
:

In this case, since  $v_I \ge v_G$ , it must be the case that  $\frac{q_G}{q_I} < \frac{(1-q)v_I(L_1+kT/r)}{qv_G(L_2-kT/r)}$ . Therefore,  $x^*=0$ , which implies that  $v_I^*=v_G^*=1$ .

So far we have assumed that  $v_I > 0$ . Suppose now  $v_I = 0$ . Then, since  $v_I \ge v_G$ , it must be the case that  $v_I = v_G = 0$ . In order that neither type of suspect deviates, it is required that  $P \le q_I(rP+T)x$  and  $P \le q_G(rP+T)x$ . Since  $x \le 1$ , this is possible, however, only if  $q_I \ge \frac{P}{rP+T}$ . Moreover, it must be that

 $x \in \left[\frac{P}{a_I(rP+T)}, 1\right]$ . Since  $v_I = v_G = 0$ , if any suspect denies, the magistrate must form an out-of-equilibrium belief  $q^*$ . The value of  $q^*$  can be arbitrary under the sequential equilibrium requirement. However, since the social cost of torture under belief  $q^*$  is  $r(1-q^*)L_1+(1-r)q^*L_2+kT$ , and the social cost of acquittal is  $q^*L_2$ , to support  $x \in \left[\frac{p}{q_I(rP+T)}, 1\right]$  as an equilibrium, it must be  $(1-q^*)L_1 + \frac{kT}{r} \le q^*L_2$ . In other words,  $q^* \in \left[\frac{L_1 + kT/r}{L_1 + L_2}, 1\right]$ .

# Appendix B: Torturing the plaintiff or the witness

We consider the two types of equilibrium in Proposition 1 separately as follows:

#### [1] Pooling equilibrium:

In this case,  $v_I^* = v_G^* = 0$  and  $x^* = 1$ , and eq. [14] becomes  $\rho Y - (1-\rho)y - \varphi T_p$ . Therefore, a plaintiff with belief  $\rho$  will come into the court if and only if  $\rho Y - (1-\rho)y - \varphi T_p \ge -\rho y$ . Note that a plaintiff with  $\rho = 0$ will never go to court, as it implies a lower utility regardless of the value of  $\varphi$ . For a plaintiff with  $\rho = 1$ , whether he will go to court depends on the value of  $Y - \varphi T_p$  versus -y. We will make the assumption that  $Y + y > T_p$ . In this case, there exists a  $\rho(\varphi) \in [0,1]$  so that a plaintiff with belief  $\rho \geq (<)\rho(\varphi)$  will (will not) come into the court, where

$$\rho(\varphi) = \frac{\varphi T_p + y}{Y + 2\nu}.$$
 [15]

Given that only those plaintiffs whose beliefs are greater than  $\rho(\varphi)$  will go to court, the belief of the magistrate that the suspect is guilty is then

$$q(\varphi) = \int_{\rho(\varphi)}^{1} \frac{\rho f(\rho)}{1 - F(\rho(\varphi))} d\rho, \qquad [16]$$

where  $F(\cdot)$  is the distribution function of  $\rho$ . Note that in our previous setup, the event starts at the time when a case has already entered into the court. We will interpret  $q(\varphi)$  as the prior of the magistrate. Since in our previous setup the magistrate does not torture the plaintiff, it corresponds to the case  $\varphi = 0$ . Consequently, q(0) = q, that is, the prior q in the previous sections is actually q(0) in the current setup.

<sup>17</sup> If  $T_p > Y + y$ , then the pain of torture is so high that a plaintiff will never come to the court.

After the plaintiff brings the case into court, the magistrate chooses an optimal probability of torturing the plaintiff,  $\varphi^*$ , to minimize the expected loss:

$$\min_{\varphi} W^{T}(\varphi) = (1 - q(\varphi))L_1 + \varphi k_p T_p.$$
 [17]

A full characterization of  $\varphi^*$  is difficult. This is because a change in  $\varphi$  will affect the magistrate's and suspect's behavior during the investigation in a complicated way that is hard to capture in a simple fashion. However, we can find a simple sufficient condition under which the magistrate will torture the plaintiff with a positive probability. Differentiating  $W^T$  with respect to  $\varphi$ , evaluated at the point where  $\varphi = 0$ , we have

$$\left.\frac{\partial W^T}{\partial \varphi}\right|_{\varphi=0} = k_p T_p - L_1 \frac{\partial q(\varphi)}{\partial \varphi}\Big|_{\varphi=0} = k_p T_p - L_1 \frac{f(\rho(0))(q-\rho(0))\rho'(0)}{1-F(\rho(0))}.$$

According to eq. [15],  $\rho'(0) = \frac{T_p}{Y+2y}$ . Thus,  $\frac{\partial W^T}{\partial \omega}|_{\omega=0} < 0$  if

$$L_1 > \frac{k_p(Y+2y)[1-F(\rho(0))]}{f(\rho(0))(q-\rho(0))} = (Y+2y)k_pH,$$
 [18]

where  $H \equiv \frac{1 - F(\rho(0))}{F(\rho(0))(q - \rho(0))}$ . This means that if  $L_1$  is sufficiently large or  $k_p$  is sufficiently small, increasing the value of  $\varphi$  slightly from 0 will decrease the value of  $W^{T}$ . Consequently, at the optimum the magistrate will torture the plaintiff with a strictly positive probability.

Inequality [18] has a clear intuition. Recall that in the pooling equilibrium all types of suspect confess, so that only the type I error is committed. If the cost of the type I error is large enough, then the benefit in terms of raising the quality of cases entering into court (by torturing the plaintiff) outweighs its cost.

Also note that eq. [18] is less likely to hold when either Y or  $\nu$  is large. This is again an intuitive result. If the plaintiff has a strong sense of "justice" in that he derives (suffers) a great utility (disutility) in having a guilty (innocent) suspect convicted, torture will be less likely to deter him from coming into court. The screening function of torture is thus small, and will be less likely to be used.

#### [2] Non-pooling equilibrium:

In this case, eq. [14] becomes

$$\rho[(1-\nu_G^*) + \nu_G^* q_G x^* r] Y - \rho \nu_G^* (1-q_G x^* r) y - (1-\rho) q_I x^* r y - \varphi T_p.$$
 [19]

As a result, he will come to the court only if  $\rho > \rho(\varphi)$ , where

$$\rho(\varphi) = \frac{\varphi T_p + q_I x^* r y}{[1 - v_G^* + v_G^* q_G x^* r](Y + y) + q_I x^* r y}.$$
 [20]

We can define the magistrate's prior  $q(\varphi)$  in a way similar to eq. [15]. Again, we assume that q(0)=q to facilitate comparison with the case when the plaintiff is not tortured. The magistrate chooses an optimal probability of torture,  $\varphi^*$ , to minimize the expected loss in this equilibrium:

$$\min_{\varphi} W^{T}(\varphi) = (1 - q(\varphi))q_{I}x^{*}rL_{1} + q(\varphi)v_{G}^{*}(1 - x^{*}q_{G}r)L_{2} 
+ [(1 - q(\varphi))q_{I} + q(\varphi)v_{G}^{*}q_{G}]x^{*}kT + \varphi k_{n}T_{n}.$$
[21]

In the case where  $v_I^* = v_G^* = 1$  and  $x^* = 0$  (i.e. region C in Figure 1), the expected payoff of the plaintiff when he comes to the court is  $-\rho y - \varphi T_p$ . As the expected payoff of a plaintiff is  $-\rho y$  if he does not come to court, we know that once  $\varphi > 0$ , no plaintiff will be coming to the court. Assuming that the magistrate's prior is q in this case, the expected loss is  $qL_2 + \varphi k_p T_p$ . Obviously, the value of  $\varphi$  that minimizes the social welfare loss is 0. Consequently,  $\varphi^* = 0$ . We thus have the following result:

**Lemma 5.** In the equilibrium in which the defendant never confesses and the magistrate never tortures the suspect, the magistrate will not torture the plaintiff either.

In region A of Figure 1 where  $x^* = v_G^* = v_I^* = 1$ , we have

$$\frac{\partial W^T}{\partial \varphi}\bigg|_{\varphi=0} = k_p T_p + \left[ (1 - q_G r) L_2 - q_I r L_1 + (q_G - q_I) k T \right] \left. \frac{\partial q(\varphi)}{\partial \varphi} \right|_{\varphi=0}, \tag{22}$$

which is negative if 18

$$L_1 + \frac{kT}{r} > \left[ \frac{q_G}{q_I} (Y + y) + y \right] k_p H + \frac{L_2}{rq_I} - \frac{q_G}{q_I} \left( L_2 - \frac{kT}{r} \right).$$
 [23]

Since  $\frac{q_G}{q_I} > \frac{(1-q)(L_1+kT/r)}{q(L_2-kT/r)}$  in region A, a sufficient condition for eq. [23] to hold is

$$L_1 + \frac{kT}{r} > q \left[ \frac{q_G}{q_I} (Y + y) + y \right] k_p H + \frac{q}{rq_I} L_2.$$
 [24]

The important thing to note is that eq. [24] is more likely to hold if  $q_G/q_I$  is small. That is, torturing the plaintiff is more likely to occur when the investigation is less informative.

**<sup>18</sup>** We use the fact that  $\frac{\partial q(\varphi)}{\partial \varphi}|_{\varphi=0} = \frac{f(\rho(0))(q-\rho(0))\rho'(0)}{1-F(\rho(0))} = \frac{\rho'(0)}{H}$ , and  $\rho'(0) = \frac{T_p}{q_Gr(Y+y)+q_Iry}$  after substituting  $x^* = \nu_G^* = \nu_I^* = 1$  into eq. [20].

In region *B* of Figure 1 where  $v_G^* \le 1$  and  $x^* \le 1$ ,

$$\frac{\partial W^{T}}{\partial \varphi}\Big|_{\varphi=0} = k_{p}T_{p} + \left\{ \frac{T}{rP+T}\underline{v}L_{2} - \frac{q_{1}P}{q_{G}(rP+T)}(rL_{1}+kT) + \frac{P}{rP+T}\underline{v}kT \right\} \frac{\partial q(\varphi)}{\partial \varphi}\Big|_{\varphi=0} + q(\varphi)[(1-x^{*}q_{G}r)L_{2} + q_{G}x^{*}kT] \frac{\partial v_{G}^{*}}{\partial q} \frac{\partial q(\varphi)}{\partial \varphi}\Big|_{\varphi=0}.$$
[25]

Since  $\frac{\partial v_G^*}{\partial q} \le 0$ , we know that  $\frac{\partial W^T}{\partial \varphi}|_{\varphi=0} < 0$  if the first two terms in eq. [25] are negative. Substituting  $\underline{v} = \frac{q_I(1-q)(L_1+kT/r)}{q_Gq(L_2-kT/r)}$  and  $\frac{\partial q(\varphi)}{\partial \varphi}|_{\varphi=0} = \frac{\rho'(0)}{H}$  into eq. [25], we know that eq. [25] is negative if

$$L_{1} + \frac{kT}{r} > \frac{q_{G}}{q_{I}} \cdot \frac{q(\frac{rP+T}{rP})(L_{2} - \frac{kT}{r})k_{p}H}{\{[q - (1-q)\frac{T}{rP}]L_{2} - \frac{kT}{r}\}\rho'(0)}.$$
 [26]

Condition [26] is more likely to hold if  $k_p$  is small, or  $\frac{\partial q}{\partial \varphi}|_{\varphi=0}$  is large (i.e. when  $\rho'(0)$  is large or when H is small). That  $k_p$  needs to be small is easy to understand. A large value of  $\frac{\partial q}{\partial \varphi}|_{\varphi=0}$  means that the average quality of cases brought into the court can be increased significantly due to a strong deterring effect of torture. Therefore, the probability of committing a type II error can be substantially reduced. Condition [26] is also more likely to hold when Y+y (which enters into eq. [26] via  $\rho'(0)$ ) is relatively small. The reason is similar to the pooling equilibrium case. If the plaintiff derives much utility from convicting the guilty, or cares much about the type II error, he will bring in cases even facing torture. In this case, the function of torture as a screening device will be small.

Again, eq. [26] is less likely to hold when  $q_G/q_I$  is large. If information revealed during investigation is relatively precise, then the chance of committing either type of error is low. In this case, there is no need to torture the plaintiff to reduce the cost of either type of error.

**Acknowledgments:** The authors thank the editor Burkhard Schipper and a referee for their helpful comments that led to significant improvements in the paper. The authors also thank Angela Ki-Che Leung for her valuable suggestions.

### References

Bodde, R., and C. Morris. 1967. Law in Imperial China: Exemplified by 190 Ch'ing (Qing) Dynasty Cases. Cambridge, MA: Harvard University Press.

Chen, K.-P., C.-F. Chou and T.-S. Tsai. 2009. "Judicial Torture as a War of Attrition." mimeo. "Confession." 2005. West's Encyclopedia of American Law. From Encyclopedia.com. Accessed December 8, 2014. http://www.encyclopedia.com/doc/1G2-3437701048.html

- Damaška, M. 1978. "The Death of Torture." The Yale Law Journal 87:860-84.
- Evans, M. D., and R. Morgan. 1998. Preventing Torture. A Study of the European Convention for the Prevention of Torture and Inhuman or Degrading Treatment or Punishment. New York: Oxford University Press.
- Fullmer, J. 1980. "Technology, Chemistry, and the Law in Early 19th-Century England." Technology and Culture 21:1–28.
- Gross, H. 1907. "Criminal Investigation: A Practical Textbook for Magistrates, Police Officers and Lawyers." In *Translated and Adapted from the System Der Kriminalistik of Hans Grossk*, edited by J. Adam and J. C. Adam. London: Specialist Press.
- Kreps, D., and R. Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27:253–79.
- Langbein, J. H. 1974. Prosecuting Crime in the Renaissance: England, Germany, France. Cambridge, MA: Harvard University Press.
- Langbein, J. H. 1977. Torture and the Law of Proof: Europe and England in the Ancien Régime. Chicago: University of Chicago Press.
- Langbein, J. H. 1978. "Torture and Plea Bargaining." *The University of Chicago Law Review* 46:3–22.
- Langbein, J. H. 1983. "'Torture." In *Encyclopedia of Crime and Justice*, edited by S. H. Kadish, 1552–6. New York: The Free Press.
- Lea, H. 1971. Superstition and Force. New York: Haskell House Publishers.
- Mialon, H., S. Mialon, and M. Stinchcombe. 2012. "Torture in Counterterrorism: Agency Incentives and Slippery Slopes." *Journal of Public Economics* 96: 33–41.
- Milgrom, P., and J. Roberts. 1982. "Predation, Reputation, and Entry Deterrence." *Journal of Economic Theory* 27:280-312.
- Nemec, J. 1968. *Highlights in Medicolegal Relations*. Bethesda, MD: National Library of Medicine.
- Parker, B. 1983. "Criminalistics." In *Encyclopedia of Crime and Justice*, edited by S. H. Kadish, 429–38. New York: The Free Press.
- Peters, E. 1985. Torture. New York: Basil Blackwell.
- Rubinstein, A. 1985. "A Bargaining Model with Incomplete Information about Time Preferences." Econometrica 53:1151–72.
- Saferstein, R. 2001. *Criminalistics. An Introduction to Forensic Science*. Upper Saddle River, NJ: Prentice Hall.
- Salop, S. C. 1979. "Strategic Entry Deterrence." *American Economic Review, Papers and Proceedings* 69:335–8.
- Sanchirico, C. W. 2000. "Games, Information, and Evidence Production: With Application to English Legal History." *American Law and Economics Review* 2:342–80.
- Sanchirico, C. W. 2001. "Relying on the Information of Interested and Potentially Dishonest Parties." *American Law and Economics Review* 3:320–57.
- Scott, G. R. 2003. The History of Torture throughout the Ages. London: Kegan Paul.
- Seidmann, D. 2005. "The Effects of a Right to Silence." *Review of Economic Studies* 72:593–614.
- Seidmann, D., and A. Stein. 2000. "The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privileges." *Harvard Law Review* 114:430–510.
- Shen, J. 1985. *Lidai Xingfa Kao (A Study on Criminal Laws throughout History*). Beijing: Zhonghua shuju.

Silverman, L. 2001. *Tortured Subjects. Pain, Truth, and the Body in Early Modern France.*Chicago: University of Chicago Press.

Spence, A. M. 1973. "Job Market Signalling." *Quarterly Journal of Economics* 87:355–74. Wantchekon, L., and A. Healy. 1999. "The 'Game' of Torture." *Journal of Conflict Resolution* 43:596–609.

Xue, Y. 1998. Tang-Ming Lue Hebian (The Tang and the Ming Codes). Beijing: Falue chubanshe.