Crime and punishment in the "American Dream" ☆

Rafael Di Tella [a,*], Juan Dubra [b]

[a] *Harvard Business School, United States*
[b] *Universidad de Montevideo, Uruguay*

Abstract

We observe that countries where belief in the "American dream"(i.e., effort pays) prevails also set harsher punishment for criminals. We know that beliefs are also correlated with several features of the economic system (taxation, social insurance, etc). Our objective is to study the joint determination of these three features (beliefs, punitiveness and economic system) in a way that replicates the observed empirical patterns. We present a model where beliefs determine the types of contracts that firms offer and whether workers exert effort. Some workers become criminals, depending on their luck in the labor market, the expected punishment, and an individual shock that we call "meanness". It is this meanness level that a penal system based on "retribution" tries to detect when deciding the severity of the punishment. We find that when initial beliefs differ, two equilibria can emerge out of identical fundamentals. In the "American" (as opposed to the "French") equilibrium, belief in the "American dream" is commonplace, workers exert effort, there are high powered contracts (and income is unequally distributed) and punishments are harsh. Economists who believe that deterrence (rather than retribution) shapes punishment can interpret the meanness parameter as pessimism about future economic opportunities and verify that two similar equilibria emerge.
© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Societies that have laws must decide what to do when people violate these laws. An interesting fact is that some societies, notably the US, choose punishments that would be considered too harsh in other societies. For example, each year in the US a small number of individuals receive capital punishment whereas in Europe killing offenders is not an option. More interesting, perhaps, is the fact that the incarceration rates in the US are extremely high. In the year 2004, for example, it exceeded 700 per 100,000 inhabitants, approximately 5 times those observed in Europe. Anecdotal

evidence suggests that even the everyday treatment of prisoners in the US is on the strict side. For example, a 1999 report states that Maricopa County Jail in Arizona makes "inmates pay for their meals, which some say are worse than those for the guard dogs. Canines eat $1.10 worth of food a day, the inmate 90 cents, the sheriff (Joe Arpaio) says".[1] Such treatment of inmates is not common in European prisons. Interestingly, crime rates for most crime categories during the early 1990's were comparable across the Atlantic (the notable exception being homicide rates), while Levitt (2004) documents a large decline — of the order of 30% — in virtually all crime categories in the US during the later part of that decade. Indeed, the criminology literature does not see the higher punitiveness of the American system as the consequence of higher crime rates. For example, Tonry (1998) opens the introductory chapter to *The Handbook of Crime and Punishment* with

> "American punitiveness is not the result of higher crime rates or of a steeper increase in crime in recent years. For most serious crimes, America's rates are not the highest among Western countries (Mayhew and van Dijk, 1997), and other countries experienced equally sharp increases in crime rates during the 1970s and 1980s (Tonry and Hatlestad, 1997, part 4). The difference is attributable to crime and punishment entanglement in American politics."[2]

Motivated by the patterns observed across Europe and America, we provide a theory of how a society's penal code and economic system are jointly determined. Recent work on "varieties of capitalism" by political scientists and economists has focused on the remarkable differences in beliefs across the Atlantic. Survey evidence reveals that the idea that "effort pays" is more prevalent in America (see, for example, Alesina and Glaeser, 2004). When asked if the poor are lazy or unlucky, over 60% of Americans while only 20% of Europeans answer that the poor are lazy rather than unlucky. Piketty (1995) exploited these differences in beliefs to show how they can give rise to two economic systems that differ greatly in the amount of government intervention even when the underlying economic processes are quite similar.[3] Given that the punitiveness of the legal system and belief in the "American dream" appear positively correlated across countries (see also Section 2 below), a natural question to ask is if these empirical patterns are predicted by models where beliefs, the penal system and the economic organization are all simultaneously determined. Note that we do not give a central role to racism, because we are interested in how people present (to themselves or to others) their views about punishment, and nobody justifies their views on account of their own racism.[4]

At least two approaches are possible in an attempt to link a belief in the "American dream" and the demand for punishment. The first is a direct link: people who believe effort pays also happen to believe that punishment should be severe. Thus an exogenous difference in beliefs gives rise to a difference in the demand for punishment. One example of such a direct link is the argument presented in Lakoff (1996) in favor of an ideological typology where right wing

---

[1] The report states that there are "chain gangs for men and women", that inmates are "forced to wear old-fashioned prison stripes and pink underwear", and that "prohibited items include cigarettes, adult magazines, hot lunches and television". This has come at considerable cost since "the county has been hit with hundreds of inmate-related lawsuits, and ordered to pay millions in legal damages". For example, in a case in April 1999, "a jury awarded .5 million to an inmate denied medical treatment for a perforated ulcer". Reported in "Arizona criminals find jail too - in 'tents'", July 27, 1999, http://www.cnn.com/US/9907/27/tough.sheriff/, accessed July 26, 2006.

[2] The relative harshness of the American penal system is also emphasized in Whitman (2003) and the literature which views punishment as a form of oppression linked to the economic order (see Wacquant (1999, 2000) for a recent example).

[3] Work by economists includes Benabou and Ok (2001), Rotemberg (2002), Benabou and Tirole (in press), Alesina and Angeletos (2005a,b), Di Tella and MacCulloch (2002), *inter alia*. A recent model in Verdier and Zenou (2004) relates beliefs about the connection between crime and race, with observed crime rates, wage premia and segregation. The political science literature is large (see, for example, the contributions in Hall and Soskice, 2001) and goes back to Alexis de Tocqueville's *Democracy in America*. Interestingly this book was written as a result of his trip to America with Gustave de Beaumont (a public prosecutor at the court of Versailles) to study the American penal system which won them the French Academy's Montyon Prize (the book was *Du systeme penitentiaire aux Etats-Unis et de son application en France*).

[4] Indeed, race and crime appear closely connected in America and the evidence documented in Glaeser and Sacerdote (2000) is certainly consistent with discrimination in sentencing. We are unaware, however, of evidence showing that states with larger proportion of black votes have also more lenient sentencing. Another simple approach is to apply Becker's model of crime. Harsh punishment in America could then be the result of higher potential bounties for criminals and/or lower legal wages relative to Europe. Alternatively, one could follow Mulligan and Shleifer (2005) and argue that setting up and running an institution requires a fixed cost so that the supply of regulation is limited by the extent of the market. Accordingly, one could argue that the administration of a serious judicial system requires a jurisdiction with a large population (such as the US relative to France). On legal justifications used for sentencing, see Tonry (1998).

individuals adhere to a "strict father" metaphor, who believe simultaneously that effort pays and that criminals should be punished.[5]

An alternative approach, and the one we follow in this paper, is to try to derive the connection between the choice of severe punishment on the one hand and the economic system induced by a belief in the "American dream" on the other.[6] In order to do this, we compare two otherwise identical societies where different beliefs about the role of luck relative to effort prevail. In our economy beliefs about how much effort pays determine the types of contracts that firms offer and whether workers accept these contracts and exert effort. The choice of becoming a criminal critically depends on a shock that we call "meanness". It is this meanness level that a penal system based on "retribution" tries to determine when establishing the severity of punishment. We find that when initial beliefs differ, two equilibria can emerge out of similar fundamentals. In the first equilibrium, which following the literature we call "American", belief in the "American dream" prevails, workers exert high effort, there are high powered contracts (there are ex-post differences in income) and punishment is harsh. In the second, "French" equilibrium, workers exert low effort and wages are independent of output. As in previous work, there is the potential for multiple equilibria (see the discussion in Section 6). If effort really pays, then American criminals are meaner than French criminals and deserve harsher punishment.

Economists who believe that deterrence (rather than retribution) shapes punishment can interpret the meanness parameter as pessimism about future economic opportunities and verify that two similar equilibria emerge. Yet another interpretation of the parameter is that of being optimistic about their future in the crime market. Indeed, if young criminals are overconfident about their future prospects, that should be modeled as a biased distribution of the parameter.[7] A focus on "retribution" in a theory of punishment, however, has several advantages (see also the discussion in Section 6). First, it is a correlate of fairness, a human tendency for which there is growing empirical evidence.[8] In this spirit, Glaeser and Sacerdote (2000) find that sentences respond to victim characteristics in a way that is hard to reconcile with optimal punishment. Second, this empirical tendency towards fairness can easily be accommodated in formal models. Models where individuals respond "like with like" and the process of detecting the amount of altruism in the other party occupies center stage (which seems particularly relevant in dispensing justice) include Levine (1998) and Rotemberg (2003). Similarly, in our theory punishment occurs because voters want to harm criminals who display low levels of benevolence towards others.[9] Third, our focus on retribution (rather than deterrence) is consistent with both practice and most work in other disciplines concerned with the problem of punishment.

Finally, our paper is related to the economics literature on crime following Becker (1968), particularly to work linking crime to redistribution (see, Harris 1970; Benoît and Osborne, 1995; Freeman, 1996, *inter alia*). For example, Benoît and Osborne (1995) analyze two different policies, and their combinations, for fighting crime. One of them is redistribution, so that the authors can also explain some of the empirical patterns that motivate our study. In their paper, however, there are exogenous differences in political processes and type of crime, whereas we explain varying levels of punishment in societies that are identical a priori.

In the next section we provide some empirical illustrations, while we present our model in Sections 3–5. Given that a model of this kind has many moving parts (at least three stages: one where labor institutions are organized, another stage where given these institutions crime decisions are taken, and a stage where sentencing of criminals is set), we keep these sections as stylized as possible. In Section 7, however, we allow a different timing which generates a richer set of empirical predictions. Importantly, it allows us to show how one group (with intermediate meanness types) can

---

[5] Benabou and Tirole (2006) explain that believers in a just world "are more likely to give stiff sentences to defendants convicted of a crime such as negligent homicide, but also to find victims (e.g., in a rape case) more culpable and "deserving" of their fate". Alesina et al. (2001) find using US data that "There is an extremely strong relationship between supporting capital punishment and opposing welfare in the US. Indeed, the correlation of these opinions ... is hardly natural. However, it makes sense if opposition to welfare comes from a desire to punish people who are seen as "stealing" from taxpayers."

[6] Nisbet and Cohen (1996), cited in Alesina et al. (2001), link punitiveness in America to the importance of the frontier, where the need to enforce uncertain property rights was salient.

[7] See Benabou and Tirole (2002) and Dubra (2004).

[8] See, for example, the growing evidence arising in the ultimatum and dictator games of Guth et al. (1982) and Forsythe et al. (1994).

[9] Related theories of punishment appear in Benoît and Dubra (2004) and Di Tella and MacCulloch (2002), although they do not have differing lengths of sentences.
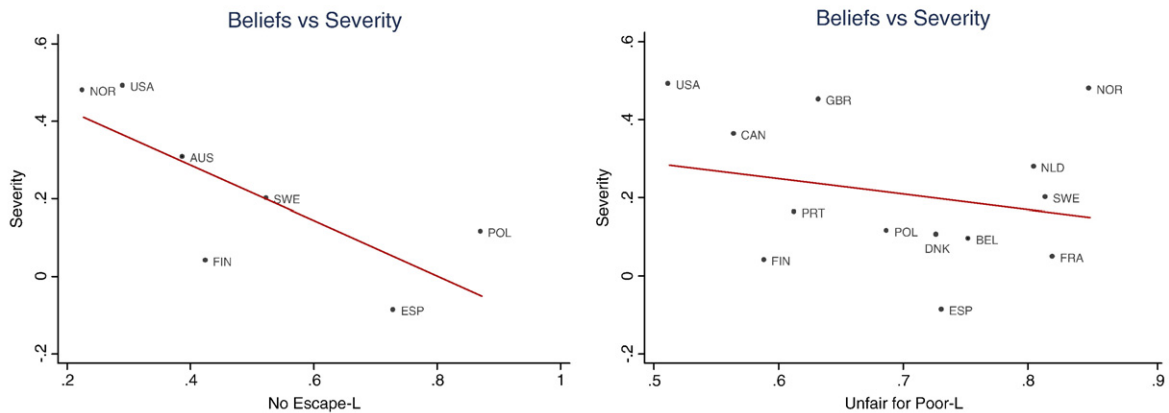
Fig. 1.

avoid crime in America by exerting effort. In that sense, crime is voluntary in America, whereas in France it is determined by luck. Section 8 concludes.

2. Beliefs, crime & punishment: empirical illustration

Considerable evidence has been gathered on the importance of beliefs about the "American dream" for the choice of economic system, including the determination of taxation, social insurance, education finance and the regulation of markets (see, Alesina and La Ferrara, 2005; Alesina et al., 2001; Fong, 2001, *inter alia*). Less work is available on the connection with desired punitiveness. A first observation is that the there are several extreme features of the American penal system such as "three-strikes-and-you-are-out laws", mandatory minimum sentence laws and "sexual psychopath" laws, which appear to be absent in other countries. The most obvious, perhaps, is the existence of the death penalty in America. This is an informal observation, however, given the well known difficulty in comparing legal systems and their punitiveness across countries (beyond the existence or not of the death penalty).[10] One exception is Alesina et al. (2001), who find that individuals who believe that effort pays also tend to be more accepting of capital punishment.

An interesting fact is the extremely high levels of incarceration in the US (interestingly, even within America, incarceration rates are higher in the 38 states that allow for the death penalty). Of course, incarceration rates are questionable as a measure of intended severity given their dependence on crime rates and enforcement efforts. Fortunately, the International Crime Victim Survey (ICVS) has asked respondents in several countries a question on desired sentencing for a similar crime.[11] The exact question employed is

> Severity: "People have different ideas about the sentences which should be given to offenders. Take for instance the case of a man of 20 years old who is found guilty of burglary for the second time. This time, he has stolen a colour TV. Which of the following sentences do you consider the most appropriate for such a case (1) Fine, (2) Prison, (3) Community service, (4) Suspended sentence, (5) Any other sentence." The variable Severity was defined as a categorical variable equals −1 if the answer is category (1), 0 if the answer is category (3), (4) or (5), and 1 if the answer is category (2).

---

[10] See Tonry (1998) and Blumstein et al. (2006). A recent US Department of Justice document comparing German and American prosecutions, states "The overall percentage of defendants convicted is also similar, but the German system has many more trials and acquittals and many fewer dismissed cases. In addition, Germany imposes sentences much shorter than those imposed in the US." See, http://www.ojp.usdoj.gov/bjs/abstract/gap.htm accessed on August 1, 2006. Difficulties in comparing severity can arise at several stages, including, pre sentencing inhibitions; the severity of the first sentence; the appeal system; the effective incarceration; and differences in treatment.

[11] For research revealing considerable consensus amongst individuals regarding the rank order severity of various offenses, see Darley et al. (1996), Kahneman et al. (1998) and Rossi et al. (1974).
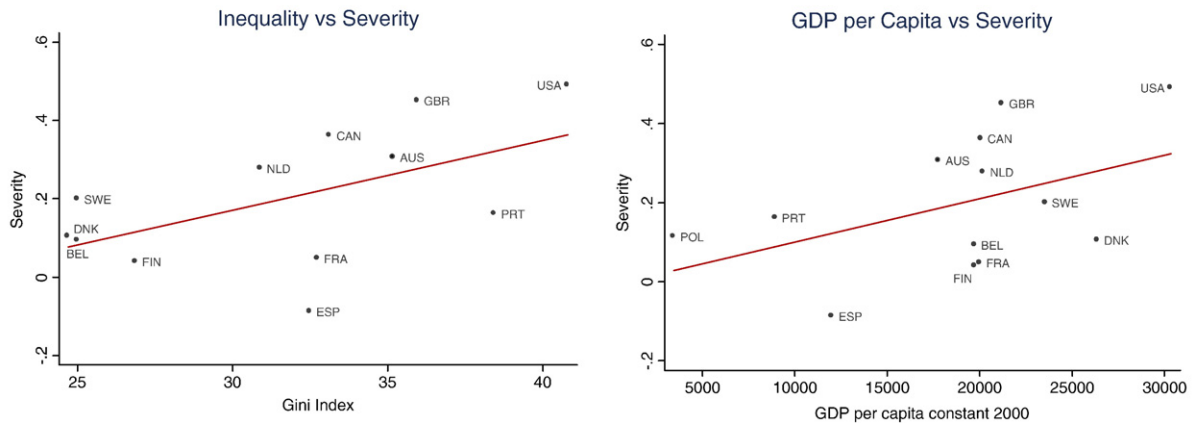
Fig. 2.

The raw data reveals that Severity is higher in the US than in Europe.[12] The only country in Europe with Severity comparable to the US is Northern Ireland (denoted NOR in the graphs below). Data on beliefs comes from four waves of the World Values Survey. We employ 2 different questions (see our working paper for further evidence). The first question is "Unfair for Poor-L": A dummy equal to 1 if the answer to the question: "Why, in your opinion, are there people in this country who live in need? Here are two opinions: which comes closest to your view? (1) They are poor because of laziness and lack of willpower, or (2) They are poor because society treats them unfairly." is (2) and 0 if the answer is (1). The second question is "No Escape-L": A dummy equal to 1 if the answer to the question: "In your opinion, do most poor people in this country have a chance of escaping from poverty, or there is very little chance of escaping? (1) They have a chance or (2) There is very little chance." was category (2) and 0 if it was category (1).

The evidence suggests that our measure of severity of punishment is positively correlated with beliefs consistent with the "American dream" (the variables are coded so that bigger numbers mean more left wing beliefs) (Fig. 1).

As a reference we note that Severity is also positively correlated with GDP per capita (measured in constant 2000 dollars) and the Gini coefficient (average 1989–2000, source: World Bank, World Development Indicators) (Fig. 2).

3. The model

We now analyze the predisposition for crime, without discussing actual crime levels (that would depend on the crime opportunities). The model is reduced form, with one firm, one worker and the Government. In the first period the firm must choose its technology and the worker must simultaneously choose his effort level. This part of our setup is a small variation on the sequential models in the literature and it is not an essential feature of the model (all we need is that in equilibrium there are multiple forms of economic organizations). In the second period the worker chooses whether to go to the crime market, and the Government chooses simultaneously the punishment level.[13]

3.1. First period

In the first period the firm can choose a market technology, in which production levels depend on the effort level of the worker, or it can choose a bureaucracy in which production is independent of effort. For $\pi_h > \pi_m > \pi_l$ (high, medium and low), profits for the firm are as follows: π_m if it chooses the bureaucracy, and π_h or π_l if it chooses the market economy and the worker chooses high or low effort respectively. The worker's total utility is his utility from the wage minus his cost of effort, which is e for high effort, and 0 for low effort. In the bureaucracy, productivity does not depend on effort, and neither does the wage which yields a utility of u_M. In the market technology output is stochastic, and the

---

[12] It is unlikely that a political economy mechanism has driven punishment away from the socially desired level. For example, Cohen et al. (2002) find that "the public largely concurs with current sentencing decisions about incarceration and sentence length". Note that actual and desired severity would be different in two countries if the apprehension rates of criminals differ.

[13] Simultaneity is not essential. All the results in the paper also obtain with sequential moves.

probability of high output is larger when the worker exerts effort. As in agency models, we assume that the firm offers high powered contracts, in which the utility received by the worker is lower when output is low. Thus, we let $\hat{u}_H$ be the utility of the worker derived from his salary when output is high, and $\hat{u}_L$ the utility of wages when it output is low: For simplicity, $\hat{u}_H$ and $\hat{u}_L$ are assumed given, as we only need that, in the optimal contract, payment in the good state is higher than in the low state. The expressions with the details of the optimal contract were used throughout a previous version of the paper. Our key results do not depend on this simplification. Payoffs are presented in the following matrix, for $p > r$.

| | | Technology choice | |
		Market	Bureaucracy
Worker	High effort	$p\hat{u}_H + (1-p)\hat{u}_L - e, \; \pi_h$	$u_M - e, \; \pi_m$
	Low effort	$r\hat{u}_H + (1-r)\hat{u}_L \quad , \; \pi_l$	$u_M \quad , \; \pi_m$

We assume that $(\hat{u}_H - \hat{u}_L)(p-r) > e$: p is high enough, relative to r; so that it compensates the cost of effort. Then, there are two equilibria in this stage of the game: a market technology with high effort and a bureaucracy with low effort (which we call the American and French equilibrium respectively). Note that two interpretations of belief in "effort pays" arise. First, effort pays to the worker because exerting effort yields a higher wage than low effort with a Market technology. This corresponds to the notion of a Nash equilibrium. Second, given a choice of Market technology by the firm, exerting effort yields a higher expected output. This notion of belief correlates with the standard view in economics of a belief as a distribution of probabilities over an unknown parameter p (the distribution is degenerate, we know with probability 1 the value of p). See Section 6 for a discussion. Our emphasis on the role of effort is consistent with the evidence in Alesina et al. (2005), who state: "hours worked per person… are almost 50% less in Europe than in the US."

The simultaneous choice of effort by the workers and of technology by the firm is not essential. For example, a similar coordination problem arises with multiple workers if in stage 1 workers invest in human capital and in stage 2 firms choose high or low skill technology (in the latter they need a few high skill workers). There would be two equilibria. One in which workers choose low effort, anticipating a choice of low skill technology by the firms, and firms then choose the low skill technology; one in which workers anticipate a choice of high skill technology, exert effort, and later firms choose the high skill technology. In order to get out of the low effort–low skill equilibrium, workers would have to coordinate and move in block towards the acquisition of more skill: that would be hard to coordinate, and is impossible within the Nash equilibrium paradigm where only single agent deviations are considered.

3.2. Second period

In the second period of the model, after his first period in the job and after observing the payoff, but before collecting it, the individual receives a shock affecting his taste for crime. The individual must then decide whether to collect his market wage, or engage in crime (discarding his wage) and the Government must choose the punishment level. Regarding the worker's decision, the interpretation is that the market wages are a first signal about his lifetime income, and he must decide whether to continue in the market, or go (forever) into the crime market.[14] If he goes honest, he collects his wages, and the taste shock is irrelevant. If he goes for crime, his payoff is $\bar{u} + \mu$, where $\bar{u}$ is the expected utility of crime and μ is the taste shock. In order to link the punishment rate with the utility of the individuals we assume that the direct utility from crime is a continuous function $\bar{u} = u(t)$ where the variable t is time in jail. Of course other dimensions (like the probability of apprehension, and the expected bounty) matter, but in order to simplify our

---

[14] There is evidence of serial correlation in wages (see for example, Baker et al, 1994). Also, we have modeled the career choice between crime and market activities as a static problem, and not one in which the individual can move back and forth between the two paths. There is some evidence that crime is an absorbing state in the sense that once the individual enters crime it is unlikely that he will get out. For example, Langan and Levin (2002) report that among nearly 300,000 prisoners released in 15 States in 1994, 67.5% were rearrested within 3 years. A prior study estimated this number at 62.5%. For an alternative timing where types are known before wages are revealed, see Section 7.

analysis, we take them as given.[15] We also assume that $\bar{u}(\cdot)$ is decreasing in t. The taste shock μ is his meanness (a large μ is a mean individual) which is drawn from a density f that is positive in its support $[\mu_{\min}; \mu_{\max}]$ and has a cumulative distribution function F.

In his decision about whether to enter the crime market, the individual compares his market utility u with $\bar{u}+\mu$ and commits a crime if and only if $u<\bar{u}+\mu$. This μ parameter can be interpreted in two ways. First, our preferred interpretation is that of a taste for crime or "meanness" shock, so that "retribution" motive for punishment is natural. Similar results can be obtained if one interprets the shock as pessimism: a large μ means that individuals perceive the economic prospects in the legal market are bleak, and are pushed into the crime market. In this case we can use a "deterrence" theory of punishment.[16]

The Government must choose the time in jail for a criminal, which in turn determines $\bar{u}$. In other words, once the Government has proved that the individual has committed a crime it must decide the time he or she must spend in jail t. The Government has a utility $v(t, \mu)$ of punishing with t years a type μ. For a belief g about the types of criminals, the Government must choose t to maximize

$$\int v(t, \mu)g(\mu)d\mu.$$

Note that prior beliefs are not related to an observable type (otherwise the process would reflect discrimination and would be invalid). We will assume that for some increasing function q, $v(t, \mu)=-(q(\mu)-t)^2$. This functional form captures the idea that if the Government knew that an individual was of type μ, it would set a punishment of $q(\mu)$, and since q is increasing, it means that the Government wants to punish "worse" individuals more. In particular, for any belief g of the government, it must choose t to maximize its utility

$$\arg\max_t -\int (q(\mu) - t)^2 g(\mu)d\mu \equiv t^* = E_g q \qquad (1)$$

This functional form, which is how we model the retributive theory of punishment, only serves to simplify our analysis (the main results do not hinge on the specific form). In the "retribution" theory of punishment that we employ, this function captures the societal desire to respond to the meanness of the criminal.[17] There is a philosophical debate that makes a distinction between the meanness of the act and the meanness of the individual who commits this act, that we do not address in this paper. We assume that retribution means giving harsher punishment to meaner individuals, as measured by a higher μ, and note that it is consistent with the practice of several legal systems (including the US and the UK). It is also consistent with the legal theories of Primoratz (1989), Kleinig (1973) and Davis (1992), who place a great deal of emphasis on the "state of mind" (mens rea) of an alleged offender, in such a way that, other things being equal, the criminal with the more malicious state of mind deserves the harsher penalty.

We note that modeling simultaneous action choices by the worker and Government is an attempt to keep "pure" our analysis of punishment as retribution. The reason is that if the Government moved first, it would take into account the effect of its punishment level on the behavior of the workers–criminals, making the theory work as a deterrence model.

4. Crime and punishment

Let g be the distribution of types that will arise in equilibrium (the distribution that the Government will presume for its calculations). We now describe the two types of equilibrium that we are interested in, which will give rise to two possible equilibrium distributions g. We are interested in two of the subgames that follow after first stage play and

---

[15] Although the probability of apprehension and time in jail can be considered partial substitutes (since both would affect deterrence of crime), it is more natural in the context of this paper to concentrate on the latter, which is chosen once the judge *knows* that the individual has committed a crime, and is thus more related to the altruism-mens rea dimensions that we explore in this paper. On this issue see the interesting paper by Mookherjee and Png (1992).

[16] See Becker (1968), Ehrlich (1975), Shavell (1987), Benoît and Osborne (1995), *inter alia*.

[17] In a "deterrence" interpretation the objective of the Government would be to minimize crime rates. Thus, the relationship between the time in jail and the $v(.)$ would be determined by the way in which criminals react to harsh sentencing.

before the worker receives his taste shock μ[18]: those following play consistent with the American and the French equilibria. We begin with a description of the game that follows after the French equilibrium which is simpler.

4.1. France

In this subgame there are two players, the worker and the Government, who must choose actions simultaneously. The action space for the worker is $C=[\mu_{min}, \mu_{max}]$: he must choose a cutoff c. The interpretation of the cutoff is that the worker will commit a crime iff $\mu \geq c$. The preceding analysis showed that for a given $\bar{u}$, the optimal cutoff is $c^*(t)=u_M-\bar{u}$, but in principle, the worker can choose any cutoff he desires. The action space for the Government is $T=[-M, M]$ for some (large) M: it must choose a time in jail t for the criminals.

The worker's expected utility for a strategy profile (c, t), given that μ has a density f, is

$$w(c,t) = \int_{\mu_{min}}^{c} u_M f(\mu)d\mu + \int_{c}^{\mu_{max}} (\bar{u}(t) + \mu)f(\mu)d\mu.$$

The first term is the payoff when μ is small and induces the worker to stay in the legal market; the second corresponds to meanness levels that induce crime. As was noted before, this payoff is maximized at $c^*(t)=u_M-\bar{u}(t)$.

Any given choice of cutoff c, not necessarily the optimal one, generates a certain criminal behavior. Conditional on having caught a criminal, and that the worker had chosen cutoff c, the Government has a posterior about the types of the criminal which is given by

$$\Pr(\mu|crime) = \frac{\Pr(cr|\mu)f(\mu)}{\Pr(crime)} = \frac{\Pr(cr|\mu)f(\mu)}{1 - F(c)} = g^c(\mu) = \begin{cases} 0 & \mu<c \\ \dfrac{f(\mu)}{1-F(c)} & \mu \geq c \end{cases}. \qquad (2)$$

To ensure continuity of the posterior beliefs we assume (Bayes does not apply) that for $c=\mu_{max}$, the Government is certain that $\mu=\mu_{max}$.

For any profile of strategies (c, t) the Government's utility is

$$G(c,t) = - \int (q(\mu) - t)^2 g^c(\mu)d\mu.$$

Then, the Government's best response to a choice of c by the worker is to set

$$t^*(c) = E_{g^c}(q) = \int q(\mu)g^c(\mu)d\mu.$$

Here we have implicitly assumed that M is large enough that $-M \leq q(\mu_{min})<q(\mu_{max}) \leq M$, so that the optimal t is feasible.[19]

The next Lemma shows one of the insights that drives many of the results in this paper: when the worker increases his cutoff, the "average" type of criminal is worse (using q and the posterior) since the best types that were formerly committing crimes are no longer criminals; this leads to a higher choice of punishment by the Government.

Lemma 1. *In the subgame following play of No Effort by the worker and Bureaucracy by the firm, the Government's best response function t^* is increasing in c.*

Proof. All proofs are in the Appendix. □

---

[18] Technically the "subgame" that follows after play of High effort-Market technology is not a subgame since in the second stage the government does not know whether the worker it faces was lucky or unlucky in the first stage. But we now analyze this second stage as a Bayesian game "as if" there had not been a first stage, but rather that the game starts by a move of nature that chooses u_H or u_L.

[19] We could have assumed that the strategy space of the Government was all of **R**; and the Government would still choose punishments in the $[q(\mu_{min}), q(\mu_{max})]$ range. We have chosen a compact strategy space so that existence of an equilibrium can be easily checked from the primitives of the model.

In a similar vein, we already know from $c^*(t) = u_M - \bar{u}(t)$ that an increase in the punishment level by the Government leads to a higher cutoff and hence to less crime by criminals that are worse on average, than before the increase in t.

Before turning to the comparative statics of the equilibrium levels of c and t with respect to the exogenous parameters, we now show that an equilibrium exists.

Proposition 1. *In the subgame following play of No Effort by the worker and Bureaucracy by the firm, there always exists an equilibrium.*

With the assumptions we have made, there may be multiple equilibria which do not affect our analysis and hence we will not make assumptions to ensure uniqueness.

We now turn to the comparative statics on u_M: how do the equilibrium values of c and t change when u_M changes? The next proposition shows that, if the equilibrium values of $(c \ (u_M), t(u_M))$ are continuous in u_M (a reasonable property of any equilibrium selection criterion and trivially satisfied when there is a unique equilibrium), either: they are increasing in u_M or the equilibrium is unstable.[20] The formal notion of stability in the following Proposition is "best case stability" of Echenique (2002).

Proposition 2. *In the subgame following play of No Effort by the worker and Bureaucracy by the firm, if the equilibrium values of c and t, as a function of u_M are continuous, they are either increasing or unstable. Moreover, there are increasing equilibrium selections.*

The result is consistent with the positive correlation between Severity and GDP per capita presented in Section 2.

4.2. America

Take a first stage play in which the individual exerts effort in his market activity and the firm chooses the market technology. Before receiving the taste shock, there will be two types of individuals: a proportion p who had a bright future because they had earned $u_H \equiv \hat{u}_H - e$ and a proportion $1-p$ who had a bleak future because they had earned $u_L \equiv \hat{u}_L - e$: The strategy of the worker is therefore a choice of two cutoffs (c_L, c_H); such that: a type who had earned u_H (a type H worker) will commit a crime if and only if $\mu \geq c_H$; a type who had earned u_L (a type L) will commit a crime if $\mu \geq c_L$. The strategy space for the worker is therefore $S = \{(c_L, c_H): c_L, c_H \in [\mu_{min}, \mu_{max}]$ and $c_L \leq c_H\}$. The Government chooses a single punishment level t, since it is not able to observe "opportunities" or first stage income.

For types H and L the utilities of cutoffs c_H and c_L when the Government chooses t, are

$$w(c, t; H) = \int_{\mu_{min}}^{c_H} u_H f(\mu) d\mu + \int_{c_H}^{\mu_{max}} (\bar{u}(t) + \mu) f(\mu) d\mu.$$

$$w(c, t; L) = \int_{\mu_{min}}^{c_L} u_L f(\mu) d\mu + \int_{c_L}^{\mu_{max}} (\bar{u}(t) + \mu) f(\mu) d\mu.$$

The best responses are given by $[c_L^*(t), c_H^*(t)] = [u_L - \bar{u}(t), u_H - \bar{u}(t)]$: The driving force behind the results in this paper, is that the Government has the same utility function as in the subgame following the French equilibrium, $v(t, \mu) = -(q(\mu) - t)^2$, but the pool of criminals it faces is different than in France and that this asymmetry was generated by the economic market outcomes. Therefore, in order to calculate the punishments by the Government, we now turn to the calculation of the Government's posterior beliefs about μ, conditional on the fact that the individual is a criminal, and that the worker had chosen a pair of cutoffs (c_L, c_H): After having caught an individual who committed a crime, the probability that he was lucky, H, in the job market is

$$P(H|\text{crime}) = \frac{P(\text{crime}|H) \frac{P(H)}{P(L)}}{P(\text{crime}|H) + P(\text{crime}|L)} = \frac{[1 - F(c_H)] \frac{p}{1-p}}{[1 - F(c_H)] + (1 - F(c_L))}$$

---

[20] The equilibrium is unstable in the sense that, starting in a value of u_M such that a slight increase in u_M does not increase $(c \ (u_M), t \ (u_M))$, it happens that the following reasoning would take you to an equilibrium with higher values of c and t (contrary to what the equilibrium selector requires): an increase in u_M makes the worker increase his cutoff for the fixed (old) equilibrium $t \ (u_M)$; given this change in the cutoff, it is optimal for the Government to increase its punishment level (by Lemma 1); given this increase in punishment, it is optimal for the worker to increase his cutoff; the process then continues with further increases in c and t, leading to a new (larger) equilibrium.

The complement is the probability that the individual was unlucky, L, given that he committed a crime: $P(L|\text{crime}) = 1 - P(H|\text{crime})$. Let $c = (c_L, c_H)$ and let f_x denote the function f on $\mu \geq x$, and 0 otherwise. We have

$$
\begin{aligned}
g^c(\mu) &= P(H|\text{crime})f(\mu|\mu > c_H) + P(L|\text{crime})f(\mu|\mu > c_L) \\[2mm]
&= P(H|\text{crime})\frac{f_{c_H}(\mu)}{1 - F(c_H)} + P(L|\text{crime})\frac{f_{c_L}(\mu)}{1 - F(c_L)} \\[2mm]
&= \frac{f_{c_H}(\mu)p + f_{c_L}(\mu)(1 - p)}{1 - pF(c_H) - (1 - p)F(c_L)} = \begin{cases} \dfrac{(1 - p)f(\mu)}{1 - pF(c_H) - (1 - p)F(c_L)} & c_L \leq \mu < c_H \\[3mm] \dfrac{f(\mu)}{1 - pF(c_H) - (1 - p)F(c_L)} & \mu \geq c_H \end{cases}
\end{aligned}
\tag{3}
$$

In order to see how our results will work, we now compare the conditional densities of types for $c_L < c_M < c_H$. We will later show that this choice of cutoffs is consistent with equilibrium choices of France and America. Notice that if $F(c_M) < pF(c_H) + (1 - p)F(c_L)$, then the comparison of Eq. (3) with the conditional density in the French equilibrium, given in Eq. (2) shows that the proportion of types larger than c_H, the really mean individuals, is larger in America than in France. The conditional density in America is such that there will be no criminal types below c_L, some types between c_L and c_H; and "a lot" of types larger than c_H. This higher proportion of mean criminals in America will then result in higher punishments.

The next Lemma presents the Government's best response to a worker's choice of criminal behavior given by $c = (c_L, c_H)$.

Lemma 2. *Let $c = (c_L, c_H)$ be a strategy choice by the worker, and let g^c be given by* Eq. (3). *The optimal punishment for belief g^c is $t^*(c; p) = E_{g^c}(q(\mu))$ or equivalently*

$$
t^*(c; p) = \int_{c_L}^{c_H} \frac{(1 - p)f(\mu)q(\mu)}{1 - pF(c_H) - (1 - p)F(c_L)}d\mu + \int_{c_H}^{\mu_{\max}} \frac{f(\mu)q(\mu)}{1 - pF(c_H) - (1 - p)F(c_L)}d\mu.
\tag{4}
$$

In order to illustrate how the model works, we now present a Lemma that will also be useful in the comparative statics that follow in the next section. It says that:

- an increase in c_L increases the Government's optimal punishment. The reason is that when c_L increases, the lowest types of μ that were committing crimes are no longer in the pool of criminals, and therefore the "average" (according to q and g^c) criminal has worsened, deserving a higher level of punishment.
- an increase in p increases the Government's optimal punishment. As p increases, the likelihood (as measured by g^c) that an individual was lucky (an H type with utility u_H, with a higher cutoff $c_H \geq c_L$) in the first period increases, and since the pool of lucky criminals is worse than the pool of unlucky criminals, the Government sets higher punishments.

Lemma 3. *The optimal punishment is increasing in c_L and in p and ambiguous in c_H.*

The proof shows that increases in p or in c_L cause a first order stochastic dominance increase in g^c (an upward movement) and, since q is increasing, its expected value increases.

Finally, we note that in this subgame an equilibrium always exists.

Proposition 3. *In the subgame following play of High Effort by the worker and Market Technology by the firm, there always exists an equilibrium.*

5. Illustration and main results

We start with a simple example that illustrates the main findings of our paper.

5.1. Illustration

The probability of high output, given high effort is 1 and given low effort 0.

		Market	Bureaucracy
	High effort	$\hat{u}_H - e,\ \pi_h$	$u_M - e,\ \pi_m$
Worker	Low effort	$\hat{u}_L\ \ \ \ ,\ \pi_l$	$u_M\ \ \ \ ,\ \pi_m$

We normalize output in France (Bureaucracy with Low Effort) to $u_M=1$ and set US (Market technology, High Effort) GDP per capita to be $\hat{u}_H=1.4$, which is consistent with observed differences in income adjusted by purchasing power parity. Also, set $e=1/10$ and $\hat{u}_L<1$, for example $\hat{u}_L=1/2$. Finally, let the expected utility of the worker when time in jail is t, be $u(t)=2-t$, and let the utility of the Government be given by $v(t,\mu)=-(q(\mu)-t)^2$ with $q(\mu)=\frac{4}{3}\mu$. The density of types μ is f, uniform on $[-2,2]$, so that $F(x)=(x+2)/4$.

We now show that in equilibrium, the optimal punishment in America is $t^A=13/5$; while in France it is $t^F=2$. When the workers know that $t^A=13/5$, they commit crimes iff

$$\mu>\hat{u}_H - e - \bar{u}(t^A) = \hat{u}_H - e - (2-t^A) = \frac{7}{5} - \frac{1}{10} - \left(2 - \frac{13}{5}\right) = \frac{19}{10}. \tag{5}$$

Since types are uniformly distributed in $[-2,2]$, $E(\mu|\text{crime})=E(\mu|\mu>19/10)=39/20$. Then, by Eq. (1), the optimal punishment is what we wanted:

$$t^A = E(q|\text{crime}) = E(4\mu/3|\text{crime}) = \frac{4}{3}E(\mu|\text{crime}) = \frac{4}{3}\frac{39}{20} = \frac{13}{5}.$$

Similarly, after the French equilibrium, when workers know that $t^F=2$, they will commit crimes if and only if $\mu>u_M-\bar{u}(t^F)=u_M-(2-t^A)=1$. This yields a conditional expectation of types (the midpoint between 1 and 2) of $E(\mu|\text{crime})=3/2$. By Eq. (1) the optimal punishment is again what we wanted.

$$t^F = E(q|\text{crime}) = E(4\mu/3|\text{crime}) = \frac{4}{3}E(\mu|\text{crime}) = \frac{4}{3}\frac{3}{2} = 2.$$

To complete the example, one only needs to check that Low Effort–Bureaucracy and High Effort–Market Technology are still equilibria of the first stage (when workers know what awaits them in the second period). This is easy and omitted.

In this example there is more punishment in America because the pool of criminals is meaner. Eq. (5) tells us that an individual will commit a crime in America if and only if his meanness is above 19/10; whereas in France he will commit a crime if and only if his meanness is larger than 1. This is the driving force behind our results.

5.2. Main results

We now return to the more general model described in Section 3, where we use high powered contracts when the firm chooses a market technology. Although this adds some complexity, it improves the explanatory power of the model (see the link between the distribution of income and punishment in Section 2).

The following is the main result of this paper. It states that there exists a density f, and a distaste for bad types q and a utility function u such that the American equilibrium leads to harsher punishments than the French equilibrium. Moreover, for any f', q' and u' close to f, q and u America still has harsher punishments. The parameters in the Proposition were chosen to match a GDP per capita that is 50% larger in the US (GDP per capita adjusted by Purchasing Power Parity is 40% larger in the US) and a poverty rate in the US of around 12% (we have chosen 10%).
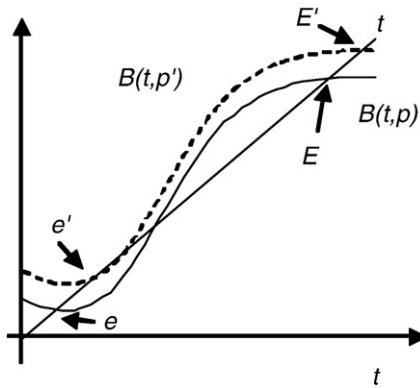
Fig. 3.

Proposition 4. *Let $u(t) = 2 - t$, $q(\mu) = \frac{4}{3}\mu$ and let f be uniform on $[-2, 2]$. With $u_M = 1$, $p = 90\%$, $r = 0\%$, $u_L = 1/2$, $u_H = 3/2$: punishment is higher in America than in France ($t^A > t^F$). Moreover, for any set of parameters and functions which are close to these, we also get the same cross country comparative statics.*

In Proposition 4 we have shown that there exists a set of parameters for which the equilibrium in America yields higher punishment rates than in France. The proof of the result shows that the lower punishment is accompanied by a higher supply of potential criminals in France. Whether this ends up in higher crime rates in France depends in the actual demand for crime (the amount for criminal opportunities), which we do not model. Note further that the evidence on relative crime rates is inconclusive (the murder rates appears higher the US, but rates for other crime categories are comparable or even higher in Europe (see Section 7). We can summarize how the Proposition works.

- There are two identical countries, which differ only in their beliefs: in America, workers believe effort pays and firms believe workers will exert effort; in France, workers believe effort doesn't pay and firms believe workers won't exert effort.
- Americans exert effort, and the French don't;
- After trying out work, individuals choose whether to stay in the labour market and collect their wages or desist and enter a life in crime. This choice depends on their income, the realization of their "meanness" shock and the punishment they would face as criminals;
- In America some good individuals who were unlucky in the labour market commit crimes. Within the group of lucky workers, only the very mean commit crimes. In France, individuals with moderate levels of "meanness" commit crimes (the kindest French criminal is worse than the kindest American criminal).[21]
- The Government punishes individuals taking the equilibrium into account, and delivering more punishment if the expected "meanness" is higher. This will result in some good but unlucky (poor) individuals also getting harsh punishment in America, but if effort really pays, these are small in number. See also the extension in Section 7, where individuals have types and some decide not to exert effort.

We now turn to an analysis of the comparative statics of the equilibrium values of the cutoffs and punishment in the American equilibrium when p and u_L change (there isn't much that can be said about changes in u_H). In order to do so, we define the composed best response function $B(t; p, u_L, u_H)$, a function of t parameterized by p, u_L and u_H, through

$$B(t; p, u_L, u_H) = t^*(c^*(t; u_L, u_H); p)$$

with $t^*(c; p)$ from Eq. (4) and $c^*(t; u_L, u_H) = [c_L^*(t), c_H^*(t)] = [u_L - \bar{u}(t), u_H - \bar{u}(t)]$.

---

[21] For an extension where individuals always know their meanness, so that those with high meanness exert no effort, see Section 7. Part of the interest in this extension is that the judge in the US faces a subset of criminals who have intermediate levels of meanness (more than the working poor but less than some of the rich) who went into a life of crime due to their unlucky draw in the labor market. This makes it worth for the US judge to investigate the background for the criminal in a search for mitigating circumstances (if they put forth effort in the past it fully reveals that they are not as mean as the other criminals).

The equilibria of the game following play of the American equilibrium in the first stage can be found by solving $B(t; p, u_L, u_H) = t$. Then, the equilibrium values of c and t are a function of the exogenous parameters p, u_L and u_H.

We begin with the comparative statics on p. This analysis will answer the question "how do equilibrium crime and punishment change when it is more true that 'effort pays'?" A similar way of interpreting the result is "what happens when there is a smaller gap between the 'American dream' and reality?"[22] Another, more subtle interpretation of the comparative statics is the following. Note that the only way in which B depends on p is through its effect on $t^*(\cdot; p)$. If instead of interpreting p as the actual probability of high output in the first period (conditional on having exerted effort) we interpret p as the belief that the Government has about the probability of high output (which could be wrong, and for example higher than the true $\tilde{p}$) the comparative statics result that follows would answer the question "how do the equilibrium crime and punishment change when the Government is run by a party that is more convinced that 'effort pays'?"

The result establishes that as people become more convinced that "effort pays", punishment becomes harsher, and crime decreases in the American equilibrium. In this game strategies are not complements, and so one can not apply the techniques of Echenique (2002) to ensure that stable equilibria have the "right" comparative statics properties (as we did in Proposition 2). With these assumptions, one can only make comparative statics assertions about the largest and smallest equilibria (or about the odd equilibria generically). Therefore, for the following results, we will assume that one of these equilibria is selected.

Proposition 5. *In the American equilibrium, when p increases, the equilibrium punishment increases and the cutoffs increase, resulting in a lower supply of potential criminals.*

Proposition 5 is illustrated in the following adaptation of Fig. 2 in Milgrom and Roberts (1994). The key for establishing our comparative statics results was showing that $B(t, p)$ is increasing in p. In particular, an increase in p keeps the worker's optimal strategy fixed (for fixed t) and makes the Government increase its optimal punishment (Lemma 3). The equilibria of this game are those in which $B(t, p) = t$:

For $p' > p$, when $B(t, p') > B(t, p)$, we get that the smallest equilibrium e moves to the right to e', and similarly for E, the largest equilibrium (Fig. 3).

Two more remarks are in order. First, as was pointed out above, our model can be interpreted as a reduced form of a principal agent model. In a in a full fledged principal–agent model the comparative statics with respect to p would be more involved, since a change in p would typically change the optimal u_H and u_L chosen by the firm. A second observation is that if one interprets the change in p not as a real change in the technology, but rather as a change in the perception of whether effort pays (say, a more right wing Government), then Proposition 5 predicts harsher punishment.

What is the effect of an increase in the wages of the poor in the American equilibrium? The next Proposition shows that it increases equilibrium punishment and reduces crime (or more precisely the supply of potential criminals). The intuition is that any given criminal is more likely to be a worse type, since there are less poor criminals and on average poor criminals have a lower meanness level μ.

Proposition 6. *In the American equilibrium, when u_L increases, the equilibrium punishment increases and the cutoffs increase, resulting in a lower supply of potential criminals.*

The proof of Proposition 6 is similar to that of Proposition 5. This result is related to a strand of the economic literature that has analyzed the link between redistribution of wealth and crime. See, Eaton and White (1991) and Benoît and Osborne (1995).

6. Discussion

6.1. Retribution and deterrence as motivation for punishment

We have assumed that retribution is the basis for punishment, although similar results obtain if we focus on deterrence. First, and beyond the three advantages listed in the introduction, retribution has long historical roots (the

---

[22] As noted above, moving p changes average income in America but it still does not mean that we are assuming a difference between the two countries because the choice of a risky technology was available to the French at the start. The French did not choose it because they did not think that effort paid.

phrase "an eye for an eye, a tooth for a tooth" is the basis for the Code of Hammurabi, one of the earliest codes of laws). Second, there is some experimental evidence that individuals are motivated by retribution concerns (over deterrence) when choosing punishment. For example, Carlsmith et al. (2002) ask respondents to read a short vignette describing a theft and the eventual apprehension of the criminal. They then ask the respondent for a sentence recommendation. The authors show that when the vignette is altered so that the probability of catching the thief changes from "almost impossible to detect" to "very easy to detect", the recommendation does not change, contrary to what deterrence suggests. They then vary the motivation of the thief (so as to change the perceived meanness of the individual). In one case he wanted money to redistribute to poorer people and in another he needed it for cancelling betting debts. This manipulation increased sentencing significantly.

Third, there are some philosophical discussions rejecting the morality of non-retributive punishment following the work of Immanuel Kant (1952), who argued that "punishment can never be administered merely as a means for promoting another good" and should be "pronounced over all criminals proportionate to their internal wickedness" (cited in Carlsmith et al., 2002). Furthermore, retribution has become the dominant theory of punishment as noted in *The Stanford Encyclopedia of Philosophy*:

> "A generation ago sociologists, criminologists, and penologists became disenchanted with the rehabilitative effects (as measured by reductions in offender recidivism) of programs conducted in prisons aimed at this end (Martinson, 1974). This disenchantment led to skepticism about the feasibility of the very aim of rehabilitation within the framework of existing penal philosophy. To these were added skepticism over the deterrent effects of punishment … and as an effective goal to pursue in punishment."

Finally, retribution is the explicit basis for some of the observed changes towards harsher punishment in America. For example, in 1976 Governor Jerry Brown signed the Determinate Sentencing Law, which made retribution the sole objective of California's sentencing system, announcing "The Legislature finds and declares that the purpose of imprisonment for crime is punishment."(California Penal Code 1170).

In order to interpret our model when the motivation for punishment is deterrence we can change the interpretation of the individual shock μ from "meanness" to "pessimism". In that case a judge facing a criminal is in front of a person who decided that his/her legal opportunities where so bad that they were seduced by a criminal life. In America there are two types. Those that faced a good initial output shock (in the proportion p) but a particularly severe pessimism shock and those that were unlucky that where on average less pessimistic than the previous group (but low initial income veered them towards crime).

A feature of crime is its high correlation with ethnic segregation, which appears higher in America. The ideas presented in the model can be extended to a setting where there are groups with different (real or perceived) net returns to education. These differences could arise because of a higher cost of getting educated or because of expected discrimination. In the extension in Section 7, criminals are punished because they are mean, although some of the rich have similar meanness (or in some cases more) than some of the criminals.

6.2. Beliefs

We model the idea that effort pays through the parameter p, so we refer to two slightly different conceptions of "effort pays". First, it suggests that if the firm chooses market technology, then effort (for the worker) pays more than being lazy. This follows the idea of a Nash equilibrium, so it really is a "belief" in the sense that if worker thinks that the firm plays *Market*, then the worker holds the belief that effort pays, which gets confirmed in equilibrium. Second, the bigger is p, the more true it is that "effort pays". This is the standard notion of a distribution of probabilities over an unknown parameter p (the distribution is degenerate in that we know with probability 1 the true value of p).[23] Note that in the American equilibrium, everyone exerts high levels of effort and a relatively large group is doing well in economic terms ($p > r$). Individuals also understand that most believe "effort pays" because

---

[23] It may be confusing that the unknown parameter p is also a probability. In this case, note that when we increase p, we can say that beliefs change (it is more true that effort pays) because the parameter increases although it would be better to say that both the truth (the true p) and beliefs about the truth have shifted towards the right (i.e., more weight on bigger values of p).

they can observe the majority of the population doing well and high powered contracts being offered. Thus, beliefs get confirmed in equilibrium.

It is worth emphasizing that we have chosen to allow for the income process (which determines whether effort pays) to be fully determined by the institutional choices of society. This is compatible with Weber's view of institutions, Denzau and North's (1994) discussion of "shared mental models", and the models where reality is "socially constructed" of Ruggie (1998) and Searle (2005). More importantly, there is no demand for new information, and there is no convergence of beliefs across the Atlantic. Even if by mistake a person wanted to change effort levels and see if their belief in the value of p is justified, such experimentation would require the cooperation of another party (i.e., it would require coordination between the firm and the worker). Contrast this with previous results by Piketty (1995), which is a single agent decision problem. There, failure to experiment results from the natural trade-off between exploration (looking for more information) and exploitation (selecting the choice that is believed to provide the highest payoff). Differences in initial beliefs lead Europe and America to different steady states. Note that full convergence on the truth is prevented by assuming an inability to observe other people's choices — the aggregate choices in the political market or the individual choices in the labor and crime market. A different approach is taken by Benabou and Tirole (2006), where a persistent desire of individuals to avoid (or distort) new information to correct a willpower problem results in lack of convergence on the "truth". In contrast, multiplicity survives in the model of Alesina and Angeletos (2005a,b), even when beliefs are unbiased, as long as the volatility of income derived from "fair" sources (savings and effort) is more sensitive to taxes than the volatility of income due to luck.[24]

6.3. The supply of potential criminals across the Atlantic

In our model, there is higher punishment in America while there is a small group of mean people predisposed to commit crime (maybe even smaller than in France). The level of crime opportunities will determine the actual crime rate. There is no obvious pattern (except for murder) across Europe and the US for most crime categories. Tonry (1998) emphasizes both a lower crime rate and a faster decline in the US; Levitt (2004) emphasizes the big decline on all crime rates in the US during the 1990's. Note also, that it is possible to argue that beliefs in effort pays are more prevalent than they should in America. Benabou and Tirole (2006) cite evidence of similar mobility levels in Europe and America. One can use this difference to obtain a high crime rate (determined by a relatively low p) and high punishment (determined by a high perception of what p is, as in Proposition 5) in America. That is, high punishment is linked to perceptions of high mobility and perceived rewards to effort, while high crime rates would follow from the reality of limited opportunities.

6.4. Link to Merton (1938)

In an influential paper, Merton (1938) argued that high crime rates in America were a result of the psychological stress created by the gap between a reality of limited opportunities and a generalized belief in the "American dream". However, Merton and subsequent research has not dealt with the problem of why punishment is so severe if such a mitigating circumstance is present. See Agnew (1999), Messner and Rosenfeld (2001), *inter alia*, as well as the review by Cullen and Agnew (2003) and the references cited therein.[25] In our model, some individuals also fail to achieve the cultural goal of success and commit crimes, although in our case this is a result of luck rather than discrimination. In our model, the prevalence of belief in the "American dream" induces a society to choose high powered incentive contracts which means that a subgroup of individuals will be poor (the unlucky). Some of them will commit crimes even though they have similar "meanness" than some rich non-criminal. It is possible to argue that the unlucky are subject to "strain" because the fact that they are more likely to commit crimes is a direct result of the prevalence of the "American dream", although it should be emphasized that we are not modelling a psychological effect that increases the predisposition to commit crimes in America.

---

[24] In Alesina and Glaeser (2004) and Di Tella and MacCulloch (2002) "unfair" income is derived from corruption, and higher taxes result in more rent seeking. Bowles and Gintis (1976) argue that status is passed on to children in part by means of unequal educational opportunities.

[25] Merton (1938) also predicts that higher crime will prevail in such societies because of the strong emphasis on success and a relatively weak emphasis on obtaining it through legitimate means. Such societies are characterized by a state of "anomie" or normlessness. See Cullen and Agnew (2003) for a description.

7. An extension where effort helps reduce criminality

Assume μ is observed before putting in effort. Again, effort pays π_h with probability p, and assume no effort pays π_l with certainty.[26] For expositional purposes assume $\bar{u}=0$.[27] Call $p\pi_h+(1-p)\pi_l-e=\mu^*$. It follows that the proportion $F(\mu^*)$ puts effort and the proportion $1-F(\mu^*)$ does not enter the labor market. Of those not entering the labor market, all are willing to commit crimes (as they obtain π_l which is lower than the return from crime, $\mu>\mu^*$). Of those working and putting forward effort, the proportion p is lucky and receives π_h and do not commit crimes (as the returns from crime are lower than the cutoff for putting no effort and are lower than the high payoff, $\mu<\mu^*<\pi_h$). The proportion $1-p$ has π_l available to them. The fraction $F(\pi_l)$ is not mean enough to commit crimes, even though their legal opportunities are bleak. The proportion $F(\mu^*)-F(\pi_l)$ commit crimes. They put forward effort but their legal opportunities were low and not sufficient to compensate them for their meanness (which, nonetheless was sufficiently low to want them to put forward effort). The model behaves otherwise similarly to the one in Sections 3–5, but we now have a richer texture of cases:

- Two otherwise identical countries differ in their beliefs about the role of effort: in one (America) workers believe effort pays and firms believe workers will work hard; in the other (France) workers believe effort doesn't pay and firms believe workers won't exert effort.
- Individuals observe their "meanness".
- In "America", some individuals (those with very large μ) decide to not work given their expected wage, and the punishment they would face in the crime market. The other group (with lower meanness) works and exerts effort. One subgroup is lucky in the labor market and earns high wages, while others are unlucky and only have low wages available to them. In France, those that decide to work do not exert effort.
- Crime decisions are made. In America, those who do not work commit crimes (Hard Criminals). Of those who work, the lucky do not commit crimes (they are Rich), while those unlucky divide into two groups, those that are relatively mean and commit crimes (the Petty Criminals) and those that are relatively good and work (the Poor). One difference between the two types of criminals is that petty criminals have some connection to the labor market (they exert effort, maybe obtaining some education or skills) while hard criminals have none. In France the mean commit crimes, while the "kind" work.
- The Government punishes taking the equilibrium into account, and delivering more punishment if the expected meanness is higher. In America the job of the judge is harder because the background of the criminals is informative of their meanness (any indication that they exerted effort is indicative of low meanness). It also has serious consequences because, if a second chance was to be given (if the game was repeated), there is some hope that the Petty Criminals would mend their ways (whereas this is not true of the Hard Criminals).
- In America, but not in France, those with intermediate meanness try to avoid a life in crime by exerting effort in the labor market. In that limited sense, crime is voluntary in America, but not in France. The following table summarizes the results of this model.

Category	Quantity	Comments
US		
Hard Criminals	$1-F(\mu^*)$	Too mean to even try. No market experience.
The Rich	$F(\mu^*)p$	Were lucky and got a high wage. But some are quite mean (the meanest in this group has type close to μ^*, the type of the kindest Hard Criminal). Some avoid crime by putting effort.
The Poor	$F(\pi_l)(1-p)$	Unlucky so legal opportunity is the low wage. They take it rather than commit crime because they are kind (have a low μ).
Petty Criminals	$(1-p)\{F(\mu^*)-F(\pi_l)\}$	Unlucky, so legal opportunity is the low wage. But they tried and put in effort. Their type is not as bad as that of the hard criminals and is comparable (sometimes better) than some of the rich, but they are meaner than the poor.
France		
Criminals	$1-F(\pi_m)$	People cannot avoid crime by exerting effort.
Honest	$F(\pi_m)$	Good people.

---

[26] This can be the result of a coordination problem or a single agent decision problem. If the latter, we would need to follow Piketty and assume that the society with lower welfare cannot observe the choices (and their results) of the other society.

[27] As long as $\bar{u}$ is not too sensitive to punishment, this is immaterial.

8. Conclusion

One unusual aspect of American social and economic organization is the harsh treatment of criminals. As Tonry (1998) describes them, "Contemporary policies concerning crime and punishment are the harshest in American history and of any western country". The objective of this paper is to provide a theory where both the supply of potential criminals and punishment as well as other features of the economic system (such as the role of incentives) are determined simultaneously in a way that fit the observed empirical patterns. In brief, we incorporate crime and punishment into an economic theory of "American exceptionalism".

As in recent work in this area, our theory gives a leading role to beliefs in the "American dream". Two otherwise identical societies start out with different degrees of belief about the impact of luck — rather than effort — in the determination of income. Two equilibria can emerge. In one society, beliefs in the "American dream" prevail, firms offer high powered incentive contracts and workers exert effort. In the second society, people tend to believe that income is independent of effort, flat contracts prevail and workers choose low effort. In our model, this occurs because of a complementarity in production whereby it does not pay for firms to offer a different type of contract given individual beliefs (although the particular mechanism is not central). Some workers become criminals, depending on their luck in the labor market, an individual shock that we call "meanness" and the expected punishment. Our main result is that punishment for criminals is harsher in the "American" equilibrium than in the "French" equilibrium. Under a "retribution" theory of punishment (as is common in the criminology literature and in practice), punishment is harsh in the "American dream" because only truly mean people would prefer to become criminals when effort really pays.[28] The model can handle an important extension, namely that types are known ex ante. In this case, different types of criminals emerge (e.g., Hard Criminals who do not even try and Petty Criminals who try but are unlucky) allowing for a richer set of empirical predictions. Given that some mean types in "America" do not commit crimes because they exerted effort and where lucky, it is possible to say (in a limited sense) that crime is voluntary. In France, in contrast, criminal careers are more the result of luck.

Appendix A. Proofs

Proof of Lemma 1. Since q is increasing it will suffice to show that the c.d.f. of $g^{c'}$ first order stochastically dominates g^c for $c' > c$:

$$\int^x g^c(\mu)d\mu = \int_c^x \frac{f(\mu)}{1 - F(c)}d\mu = \frac{F(x) - F(c)}{1 - F(c)} \geq \frac{F(x) - F(c')}{1 - F(c')} = \int^x g^{c'}(\mu)d\mu. \quad \square$$

Proof of Proposition 1. Continuity of the utility functions follows from continuity of $\bar{u}$ and because f is a density, ensuring continuity of its posterior g^c. The strategy spaces are compact, so quasiconcavity of the utilities in own strategy establish the result. The derivative of $w(c, t)$ wrt c is $[u_M - \bar{u}(t) - c] f(c)$ which is positive for small values of c and then negative; and the derivative of $G(c, t)$ with respect to t is $\int 2 (q(\mu) - t) g^c (\mu) d\mu$ which is positive and then negative, proving quasiconcavity. $\quad \square$

Proof of Proposition 2. Increasing equilibria exist by Corollary 1 in Milgrom and Roberts (1994) applied to $t^*(c^*(t; u_M))$. Non-increasing continuous equilibrium selections are unstable because Lemma 1 shows that $t^*(c)$ is increasing, and from $c^*(t; u_M) = u_M - \bar{u}(t)$ we obtain that $c^*(t; u_M)$ is also increasing, so that $[c^*(t; u_M), t^*(c)]$ is an increasing family of functions as required by Theorem 2 in Echenique (2002) which then yields our desired result. $\quad \square$

---

[28] When punishment is set for "deterrence" reasons (as is common in economic models), harsh punishment corrects (ex ante) the tendency to commit crimes of the pessimists, which are over represented in the criminal population in the equilibrium with high powered contracts (relative to "France").

Proof of Lemma 2. The FOC for the maximization problem of the Government is

$$\int \frac{\partial v(t,\mu)}{\partial t} g^c(\mu) d\mu = \int 2(q(\mu) - t) g^c(\mu) d\mu = 0 \Leftrightarrow t^*(c;p) = E_{g^c}(q(\mu)). \qquad \square$$

Proof of Lemma 3. We must show that the c.d.f. G^c of g^c increases in first order stochastic dominance sense when p or c_L increase. For $p' > p$ we have that

$$
G^c(x;p) = \int_{c_L}^x \frac{f(\mu)(1-p)}{1 - pF(c_H) - (1-p)F(c_L)} d\mu + \int_{c_H}^x \frac{pf(\mu)}{1 - pF(c_H) - (1-p)F(c_L)} d\mu
$$

$$
= \begin{cases}
\dfrac{(F(x) - F(c_L))(1-p)}{1 - F(c_H)p - F(c_L)(1-p)} & x \le c_H \\[2ex]
\dfrac{F(x) - F(c_H)p - F(c_L)(1-p)}{1 - pF(c_H) - (1-p)F(c_L)} & x > c_H
\end{cases}
\tag{6}
$$

The derivative of G^c with respect to p is then

$$
\frac{dG^c(x;p)}{dp} = \begin{cases}
-\dfrac{(F(x) - F(c_L))(1 - F(c_H))}{[1 - pF(c_H) - (1-p)F(c_L)]^2} \le 0 & x \le c_H \\[2ex]
\dfrac{(F(c_L) - F(c_H))(1 - F(x))}{1 - pF(s_H) - (1-p)F(c_L)]^2} \le 0 & x > c_H
\end{cases}
$$

We have therefore shown that for all x, $G^c(x; p') \le G^c(x; p)$ for $p' > p$, as was to be shown.

Now set $q'_L > c_L$ and note that we get again a fosd increase in G^c; completing the proof

$$
\frac{dG^c}{dc_L} = \begin{cases}
f(c_L)(1-p)\dfrac{F(x)(1-p) + F(c_H)p - 1}{[1 - pF(c_H) - (1-p)F(c_L)]^2} \le 0 & x \le c_H \\[2ex]
\dfrac{f(c_L)(1-p)(F(x) - 1)}{[1 - pF(c_H) - (1-p)F(c_L)]^2} \le 0 & x > c_H
\end{cases}.
$$

The ambiguity in d_t/dc_H obtains because for some densities f the posteriors for the two levels of c_H won't be ranked in fosd, and therefore, for some qs the low level of c_H yields higher punishment, and for some qs the high level of c_H will yield higher punishments. $\square$

Proof of Proposition 3. Given our assumptions, the aggregate best response $B(c, t) = [[c_L^*(t), c_H^*(t)], t^*(c)]$ is a continuous function in a convex and compact subset of the $\mathbf{R}^3$, so that by Brouwer's theorem, it has a fixed point, which is an equilibrium. $\square$

Proof of Proposition 4. America. Following choices of a market technology and high effort in the first period, we have that

$$
g^A = \begin{cases}
\dfrac{1-p}{4 - t - pu_H - (1-p)u_L} & u_H + t - 2 \ge \mu > u_L + t - 2 \\[2ex]
\dfrac{1}{4 - t - pu_H - (1-p)u_L} & 2 \ge \mu > u_H + t - 2
\end{cases}.
$$

Then, using

$$E_{g^A}\mu = P(\mu \le u_H + t - 2)E(\mu|\mu \le u_H + t - 2) + P(\mu > u_H + t - 2)E(\mu|\mu > u_H + t - 2)$$

and that conditional expectations are the midpoints of the intervals, we get

$$\frac{3t}{4} = E_{g^A}\mu = \frac{(1-p)(u_H - u_L)}{4 - t - pu_H - (1-p)u_L}\left(\frac{u_H + u_L}{2} + t - 2\right) + \frac{4 - u_H - t}{4 - t - pu_H - (1-p)u_L}\frac{u_H + t}{2}.$$

The solution to the previous equation then yields the equilibrium punishment

$$
t^A = \frac{\dfrac{4 + pu_H + (1-p)u_L}{2}}{-\dfrac{\sqrt{-24(pu_H + (1-p)u_L) + 2pu_H u_L + 9u_L^2 + 8pu_H^2 - 10pu_L^2 + p^2(u_H - u_L)^2 + 16}}{2}}
\tag{7}
$$

Equilibrium, crime is then given by

$$
1 - pF(u_H - \bar{u}) - (1-p)F(u_L - \bar{u}) = 1 - pF(u_H + t - 2) - (1-p)F(u_L + t - 2)
$$
$$
= \frac{4 - t - pu_H - (1-p)u_L}{4}.
\tag{8}
$$

France. Punishment in France is given by the density in Eq. (2), $g^F(\mu) = \frac{1}{4 - u_M - t}$ for $\mu \geq u_M + t - 2$; so that the expected value of μ is

$$
E_{g^F}\mu = \frac{u_M + t}{2} \Leftrightarrow \frac{3t}{4} = \frac{u_M + t}{2} \Leftrightarrow t^F = 2u_M.
\tag{9}
$$

Crime in France will then be

$$
\Pr(\mu \geq u_M - \bar{u}) = \Pr(\mu \geq u_M + t - 2) = \frac{4 - u_M - t}{4} = \frac{4 - 3u_M}{4}.
\tag{10}
$$

We will show that punishment is higher in America and crime higher in France. From Eqs. (7) and (9) we get $t^A = 2.26 > 2 = t^F$, as was to be shown. Notice that this implies sentencing standards 13% higher in the US. Moreover, from Eqs. (8) and (10) we see that crime in the US is 8% whereas in France it is 25%.

Finally, we must check that the two first period equilibria are indeed equilibria. We let the probability of high output, with low effort, in the market technology be 0; and let e, the cost of effort, be arbitrarily small. Suppose then that the firm chooses a market technology. The expected payoff for the individual is then the combination of whether he will receive u_H or u_L, and what will happen after observing μ:

$$
p\left[F(u_H - \bar{u})u_H + \int_{u_H - \bar{u}}^{2} (\bar{u} + \mu)f(\mu)d\mu\right] + (1-p)\left[F(u_L - \bar{u})u_L + \int_{u_L - \bar{u}}^{2} (\bar{u} + \mu)f(\mu)d\mu\right].
$$

This equation becomes 1:4254; from the following calculation (with $z = (22 - \sqrt{19})/10$):

$$
\frac{9}{10}\left[\frac{126 - 3\sqrt{19}}{80} + \int_{z}^{2}\left(\frac{\sqrt{19} - 7}{40} + \frac{\mu}{4}\right)d\mu\right] + \frac{1}{10}\left[\frac{32 - \sqrt{19}}{80} + \int_{z}^{2}\left(\frac{\sqrt{19} - 7}{40} + \frac{\mu}{4}\right)d\mu\right]
$$

Notice that the ex ante expected value is larger than 1:4, which is $pu_H + (1-p)u_L$, given that the individual can re-optimize after observing u_H or u_L. The uncertainty in the utility of playing low effort comes from the unknown μ We then have that because the probability of high output is 0 and the cost of effort is almost 0, the expected utility of the worker of choosing low effort when the firm chooses the market technology is

$$
F(u_L - \bar{u})u_L + \int_{u_L - \bar{u}}^{2} (\bar{u} + \mu)f(\mu)d\mu = \frac{32 - \sqrt{19}}{80} + \int_{\frac{12 - \sqrt{19}}{10}}^{2}\left(\frac{\sqrt{19} - 7}{10} + \mu\right)\frac{1}{4}d\mu = 0.69093.
$$

We conclude that if the firm chooses the market technology it is better to choose high effort.

Suppose now that the workers are choosing high effort. Then, the firm must choose between π_h and π_m, and we had assumed $\pi_h > \pi_m$, so indeed the American equilibrium is an equilibrium.

We now check that the French equilibrium is an equilibrium. First notice that if the workers are choosing low effort, the firm would rather choose a bureaucratic technology, because $\pi_m > \pi_l$. If the firm chooses a bureaucracy, for each value of μ the individual is better off choosing low effort, and for some values, strictly better off because of the cost of effort. Hence, a bureaucracy with low effort is also an equilibrium.

To see that parameters and functions close to the ones we have chosen also have the property that $t^A > t^F$, note that $t^A = E_{g^A}q$ and $t^F = E_{g^F}q$ are continuous functions of all parameters and functions involved. Similarly, crime rates are a continuous function of all the parameters and functions, and of the punishment rate. Since small changes in parameters and functions lead to slight changes in punishment, crime will also change slightly. $\square$

Proof of Proposition 5. We will show that B $(t; p, u_L, u_H)$ is continuous in t and increasing in p, and then apply Corollary 1 in Milgrom and Roberts (1994). Continuity follows from: the continuity of $(u_L - \bar{u}(t), u_H - \bar{u}(t))$ in $\bar{u}(t)$; the continuity of $\bar{u}(t)$ in t; and the continuity of $t^*(c; p)$ when the distribution of types μ has a density (as we have assumed). To show that B is increasing in p, notice that when p increases, for a fixed t; $c^*(t; u_L, u_H)$ is unchanged. Then, for any fixed pair of cutoffs (c_L, c_H) (in this case, a fixed $c^*(t; u_L, u_H)$) of the worker, the increase in p causes an increase (in first order stochastic dominance sense) of the posterior belief in the Government's posterior and therefore in the Government's optimal punishment $t^*(c^*(t; u_L, u_H); p)$ as was shown in Lemma 3. $\square$

Proof of Proposition 6. The structure of the proof is as that of Proposition 5. Continuity has already been shown, so we will only show that $D(t; u_H, u_L, p) \equiv B^g(B^w(t))$ is increasing in u_L. For a fixed t and $u'_L > u_L$, we get $B^w(t; u'_L) = (u'_L - \bar{u}(t), u_H - \bar{u}(t))$ and $B^w(t; u_L) = (u_L - \bar{u}(t), u_H - \bar{u}(t))$, so it will suffice to show that for $s'_1 > s_1$ we obtain $B^g((s'_1, s_2)) \geq B^g((s_1, s_2))$. This follows from the fact that the posterior of f for (s'_1, s_2) first order stochastically dominates the posterior of f for (s_1, s_2): for G^A as in Eq. (6) we get what we wanted to show

$$
\frac{dG^A}{ds_1} = \begin{cases} f(s_1)(1-p)\dfrac{F(x)(1-p) + F(s_2)p - 1}{[1 - F(s_2)p - F(s_1)(1-p)]^2} \leq 0 & x \leq s_2 \\[4mm] \dfrac{f(s_1)(1-p)(F(x)-1)}{[1 - F(s_2)p - F(s_1)(1-p)]^2} \leq 0 & x > s_2 \end{cases} \qquad \square
$$

References

Agnew, Robert, 1999. A general strain theory of community differences in crime rates. Journal of Research in Crime and Delinquency 36, 123–155.

Alesina, Alberto, Glaeser, Ed, 2004. Fighting Poverty in the US and Europe: A World of Difference. Oxford University Press, Oxford UK.

Alesina, Alberto, Angeletos, George-Marios, 2005a. Corruption, inequality and fairness. Journal of Monetary Economics 52 (7), 1227–1244.

Alesina, Alberto, Angeletos, George-Marios, 2005b. Fairness and redistribution. American Economic Review 95 (4), 960–980.

Alesina, Alberto, La Ferrara, Eliana, 2005. Redistribution in the land of opportunities. Journal of Public Economics LXXXIX, 897−931.

Alesina, Alberto, Glaeser, Edward, Sacerdote, Bruce, 2001. Why doesn't the US have a European style welfare state? Brookings Papers on Economic Activity.

Alesina, Alberto, Glaeser Edward, Sacerdote, Bruce, 2005. Work and Leisure in the US and Europe: Why so Different? Mimeo.

Baker, George, Gibbs, Michael, Holmstrom, Bengt, 1994. The internal economics of the firm: evidence from personnel data. Quarterly Journal of Economics 109.

Becker, G., 1968. Crime and punishment: an economic approach. Journal of Political Economy 76 (2), 169–217.

Benabou, Roland, Ok, Efe A., 2001. Social mobility and the demand for redistribution: the POUM hypothesis. Quarterly Journal of Economics 116 (2), 447–487.

Benabou, Roland, Tirole, Jean, 2002. Self-confidence and personal motivation. Quarterly Journal of Economics 117 (3), 871–915.

Benabou, Roland, Tirole, Jean, 2006. Belief in a Just World and Redistributive Politics. Quarterly Journal of Economics 121 (2), 699–746.

Benabou, Roland, Tirole, Jean, in press. Belief in a Just World and Redistributive Policies, forthcoming. Quarterly Journal of Economics.

Benoît, J.-P., Osborne, M., 1995. Crime, punishment, and social expenditure. Journal of Institutional and Theoretical Economics 151, 326–347.

Benoît, J.-P., Dubra, J., 2004. Why do good cops defend bad cops? International Economic Review 45.

Blumstein, A., Tonry, M., van Ness, A., 2006. Cross-national measures of punitiveness. In: Tonry, Michael, Farrington, David P. (Eds.), Crime and Punishment in Western Countries, 1980–1999. Chicago University Press.

Bowles, Samuel, Gintis, Herbert, 1976. Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. Basic Books, New York.

Carlsmith, K.M., Darley, J.M., Robinson, P.H., 2002. Why do we punish? Deterrence and just desserts as motives for punishment. Journal of Personality and Social Psychology 83, 284–299.

Cohen, Mark, Rust, Roland, Steen, Sara, 2002. Measuring Public Perceptions of Appropriate Prison Sentencesls. Mimeo.

Cullen, Francis, Agnew, Robert, 2003. Anomie/strain theories of crime, Part V. In: Cullen, Francis, Agnew, Robert (Eds.), Criminological Theory: Essential Readings. Roxbury Publishing Company, Los Angeles, California.

Darley, J.M., Sanderson, C.A., LaMantia, P.S., 1996. Community standards for defining attempt: inconsistencies with the model penal code. American Behavioral Scientist 39, 405−420.

Davis, Michael, 1992. To Make the Punishment Fit the Crime, Essays in the Theory of Criminal Justice. Westview Press.

Denzau, Arthur, North, Douglass, 1994. Shared mental models: ideologies and institutions. Kyklos 47 (1), 3–31.

Di Tella, Rafael, MacCulloch, Robert, 2002. Why doesn't Capitalism Flow to poor Countries? Mimeo. Harvard Business School.

Dubra, J., 2004. Optimism and overconfidence in search. Review of Economic Dynamics 7 (1), 198–218 2004.

Eaton, B.C., White, W.D., 1991. The distribution of wealth and the efficiency of institutions. Economic Inquiry 29, 336–350.

Echenique, F., 2002. Comparative statics by adaptive dynamics and the correspondence principle. Econometrica 70 (2).

Ehrlich, I., 1975. The deterrent effect of capital punishment: a question of life and death. American Economic Review 55, 397–417.

Fong, Christina, 2001. Social preferences, self interest, and the demand for redistribution. Journal of Public Economics LXXXII, 225–246.

Forsythe, R., Horowitz, J.L., Savin,, N.E., Sefton, M., 1994. Fairness in simple bargaining games. Games and Economic Behavior 6, 347–369.

Freeman, Richard, 1996. Why do so many young American men commit crimes and what might we do about it? Journal of Economic Perspectives 10 (1), 25–42.

Glaeser, Ed, Sacerdote, Bruce, 2000. The determinants of punishment: deterrence, incapacitation and vengeance. NBER Working Paper, vol. W7676.

Guth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. Journal of Economic Behavior and Organization 3, 367–388.

Hall, Peter, Soskice, David, 2001. Varieties of Capitalism: The Institutional Foundations of Comparative Advantage. Oxford University Press, Oxford.

Harris, J., 1970. On the economics of law and order. Journal of Political Economy 78 (1), 165–174.

Kant, Immanuel, 1952. The science of right (W. Hastie, Trans.). In: Hutchins, R. (Ed.), Great books of the Western world, vol. 42. Kant (pp. 397−446). Chicago: Encyclopedia Britanica. Hillsdale, NJ: Erlbaum.

Kleinig, John, 1973. Punishment and Desert. Martinus Nijhoff, The Hague.

Kahneman, D., Schkade, D., Sunstein, C.R., 1998. Shared outrage and erratic awards: the psychology of punitive damages. Journal of Risk and Uncertainty 16, 49−86.

Lakoff, George, 1996. Moral Politics: How Liberals and Conservatives Think. University of Chicago Press, London.

Langan, Patrick, Levin, David, 2002. Recidivism of Prisoners Released in 1994. Bureau of Justice Statistics Special Report.

Levine, David K., 1998. Modelling altruism and spitefulness in experiments. Review of Economic Dynamics 1, 593–622.

Levitt, Steven, 2004. Understanding why crime fell in the 1990s: four factors that explain the decline and six that do not. Journal of Economic Perspectives 18 (1).

Nisbet, R., Cohen, D., 1996. Culture of Honor: The Psychology of Violence in the South. Westview Press, Boulder.

Martinson, Robert, 1974. What works? Questions and answers about prison reform. Public Interest 10, 22–54.

Mayhew, Pat, van Dijk, Jan, 1997. Criminal Victimization in 11 Industrialized Countries. Dutch Ministry of Justice, The Hague.

Merton, Robert K., 1938. Social structure and anomie. American Sociological Review 3, 672–682 October.

Messner, Steven, Rosenfeld, Richard, 2001. Crime and the American dream, Chapter 19. In: Cullen, Francis, Agnew, Robert (Eds.), Criminological Theory: Essential Readings. Roxbury Publishing Company, Los Angeles, California.

Milgrom, P., Roberts, J., 1994. Comparing equilibria. American Economic Review 84 (3), 441–459.

Mookherjee, Dilip, Png, Ivan, 1992. Monitoring vis a vis investigation in the enforcement of law. American Economic Review 82 (3), 556–565.

Mulligan, Casey, Shleifer, Andrei, 2005. The extent of the market and the supply of regulation. Quarterly Journal of Economics 120 (4), 1445–1473.

Piketty, Thomas, 1995. Social mobility and redistributive politics. Quarterly Journal of Economics 110, 551–584.

Primoratz, Igor, 1989. Justifying Legal Punishment. Humanities Press International, Atlantic Highlands, NJ.

Rotemberg, Julio J., 2002. Perceptions of equity and the distribution of income. Journal of Labor Economics XX, 249−288.

Rotemberg, Julio, 2003. Minimally Acceptable Altruismand the Ultimatum Game. Mimeo.

Rossi, P.H., Waite, E., Bose, C.E., Berk, R.E., 1974. The seriousness of crimes: normative structure and individual differences. American Sociological Review 39, 224−237.

Ruggie, John, 1998. What makes the world hang together? Neo-utilitarianism and the social constructivist challenge. International Organization 52 (4), 855–885.

Searle, John, 2005. What is an institution? Journal of Institutional Economics 1 (1).

Shavell, Steven, 1987. The optimal use of nonmonetary sanctions as a deterrent. American Economic Review 77 (4), 584–592.

Tonry, Michael, 1998. Introduction: crime and punishment in America. In: Tonry, Michel (Ed.), The Handbook of Crime and Punishment. Oxford University Press, New York.

Tonry, Michael, Hatlestad, Kathleen, 1997. Sentencing Reform in Overcrowded Times, a Comparative Perspective. Oxford University press, New York.

Verdier, Thierry, Zenou, Yves, 2004. Racial beliefs, location and the causes of crime. International Economic Review 45 (3), 731−60.

Wacquant, Loïc, 1999. Le Prisions de la misêre. Éditions Raisons D'Agir, Ed Manantial.

Wacquant, Loïc, 2000. Deadly symbiosis: when ghetto and prison meet and mesh. Punishment and Society 3 (1), 95–134.

Whitman, James Q., 2003. Harsh Justice: Criminal Punishment and the Widening Divide between America and Europe. Oxford University Press, Oxford.