CRIME, PUNISHMENT AND SOCIAL NORMS

JÖRGEN W. WEIBULL AND EDGAR VILLA

STOCKHOLM SCHOOL OF ECONOMICS AND BOSTON UNIVERSITY

SSE/EFI WORKING PAPER SERIES IN ECONOMICS AND FINANCE NO 610.

November 10, 2005

ABSTRACT. We analyze the interplay between economic incentives and social norms when individuals decide whether or not to engage in criminal activity. More specifically, we assume that there is a social norm against criminal activity and that deviations from the norm result in feelings of guilt or shame. The intensity of these feelings is here endogenous in the sense that they are stronger when the population fraction obeying the norm is larger. As a consequence, a gradual reduction of the sanctions against criminal activity, or of the taxation of legal incomes, may weaken the social norm against crime. Due to the potential multiplicity of equilibria in our model, such a gradual change may even induce a discontinuous increase in the crime rate. We show that law enforcement policies may have dramatic and permanent effects on the crime rate, and lead to hysteresis. We also define political equilibrium under majority rule and show how a majority of individuals, who feel no guilt or shame from violating the law, in political equilibrium can exploit a minority who do have such feelings.

Keywords: crime, punishment, social norm, political equilibrium.
JEL-codes: D01, D11, D72.

1. INTRODUCTION

Human behavior seems to be influenced both by economic incentives and social norms. While sociologists have emphasized social norms since the time of Emile Durkheim, economists have focused almost exclusively on economic incentives since the time of Adam Smith. Most likely, the different approaches reflect the different subject matters of the two disciplines; while sociologist mostly study behavior in social groups, economists focus mainly on behavior in the marketplace. However, for some decisions that people make, both social norms and individual economic incentives appear to be involved. This paper is an attempt to bring together social norms and individual rational choice in an economic model of crime following the seminal model of Becker

(1968). More precisely, we extend the traditional economic approach, where pecuniary motives are the driving force of the criminal offenses, to encompass social norms in such a way as to enable an analysis of the interplay between economic incentives and social norms. The treatment of social norms is similar to that in Lindbeck-Nyberg-Weibull (1999) and Huck-Kübler-Weibull (2004) in the way social norms are modeled. The first of those studies analyzes the decision whether to work or live off transfers, and the second analyzes individual effort decisions in team work. By contrast, we here analyze individual decisions of whether or not to take part in criminal activity.

On top of the legal sanctions taken against criminals, there seems to exist in most societies an implicit social norm to live off legal activities rather than illegal. This normative pressure can be more or less strong, and can take the form of "shame" or "guilt". As Elster (1989) has pointed out "if punishment was merely the price tag attached to crime, nobody would feel shame when caught" p. 105. By shame is usually meant embarrassment in front of others who have observed or know about the deviation, while guilt refers to internalized shame, or embarrassment in front of oneself. We believe that both shame and guilt are at work in connection with criminal behavior.

We also believe that the intensity of the social norm to stay out of crime depends in part on how many others in one's peer group stay out of crime. The more frequent crime is, the lesser the guilt and shame attached to it. Bentham (1789) referred to this effect as the syndrome of robberies without shame: "where robberies are frequent, and unpunished, robberies are committed without shame" p. 156. Living off criminal activity is less shameful the more criminal activity there is in society. While the existence of a social norm against crime is taken as a given here, the intensity of it, as perceived by the individual, is endogenous in the model: it depends positively on the population fraction adhering to it. Hence, the intensity of the social norm becomes an equilibrium phenomenon, which, *a priori*, allows for the possibility of multiple equilibria under given taxes and sanctions against crime.

We develop a relatively simple model and study in detail two cases. In the first case, the return to criminal activity is uncorrelated with the individual's potential wage from legal work. This may, for example, be the case with illegal dealing with drugs or weapons, auto-theft and robbery. In the second case, the returns to criminal activity are positively correlated with potential wages: the higher an individual's potential wage from legal work, the higher are the expected returns from criminal activity. This may plausibly be the case with tax evasion, economic extortions and theft within firms and organizations. We assume, throughout, that every individual has only a binary choice: either to work off legal or illegal activity. The individual obtains material and immaterial utility from his or her choice, where the material utility depends on the disposable income from legal work as well as on the sanctions against crime and the probability of being caught. The immaterial utility emanates from norm adherence. If driven only by guilt, this is independent of the probability of being caught and convicted, while if driven by shame it does depend on this probability.

Within this modelling framework we first analyze individual decision-making and equilibrium outcomes under given policy parameters concerning taxation of legal work and law enforcement measures. Each individual in the economy then decides whether or not to engage in criminal activity, and these decisions are based on knowledge of relevant policy parameters and on expectations about the going crime rate. The latter is important, because it may influence the utility from norm adherence. A crime rate is an equilibrium outcome if no individual wants to change his or her individual choice at that crime rate. Because of the endogeneity of the intensity of the social norm, there may be multiple equilibria for given policy parameters. We establish a sufficient condition for the existence of equilibrium and another sufficient condition for its uniqueness. We also illustrate the equilibrium outcome and its comparative statics properties by means of examples. Having studied equilibrium crime rates, we turn to the question of how policy is determined. We call a policy p, such as a combination a tax rate and a penalty for criminals, an equilibrium policy if (i) it is consistent with equilibrium behavior and a balanced government budget, and (ii) there exists no other such policy, p', that would gain a strict majority against p if individuals would vote earnestly according to the expected utility they obtain under the respective policy. We illustrate the nature of equilibrium policies within the context of a tax evasion example.

In order to high-light the interplay between economic incentives and social norms we make several simplifying assumptions. One such assumption is that we take the probability of catching a given criminal as independent of the current crime rate. This is clearly unrealistic if the capacity of the law enforcement system is fixed and given. However, our assumption makes the analysis and exposition easier, and, once this case has been treated, it is not difficult to extend the present analysis to the case of an endogenous such probability — another source of potential equilibrium multiplicity.[1]

[1]Such a generalization can be readily obtained by letting the punishment probability π be a continuous and non-increasing function of the crime rate x, see section 2. See also section 4.1 where

The rest of the paper is organized as follows. The model is developed in section 2, equilibrium crime rates are analyzed in section 3 and section 4 is devoted to political equilibrium considerations.

2. The model

Consider a continuum population of individuals, where each individual faces a binary choice: whether or not to engage in some criminal activity. Let x denote the population share of individuals who choose *crime*. An individual who does not choose to engage in crime will be said to choose *work*, but we do not exclude the possibility that criminals also work.

We assume the labor market to be perfectly competitive: each worker is paid his or her marginal productivity or *potential wage,* w, which is exogenous and fixed. Labor incomes are taxed at a constant rate τ. The disposable (life time) income to an individual with potential wage w who chooses *work* is $(1-\tau)\,w$. The disposable income to the same individual, when instead choosing *crime*, is $\sigma_1 w+\mu_1$ if *not* caught and convicted and otherwise $\sigma_0 w+\mu_0$. We assume

$$\sigma_0 w+\mu_0<\sigma_1 w+\mu_1, \tag{1}$$

that is, the disposable income of a criminal who is caught and convicted is lower than that of a criminal who is not. The σ-parameters are non-negative and account for a positive "correlation" between a criminal's earnings, or disposable income, and his or her potential wage. This can be due to full-time or part-time legal work. Moreover, higher paid employee's within a firm's hierarchy usually have more control of and knowledge about the firm's resources, thereby opening up the possibility for higher criminal earnings, all of which suggests $\sigma_1>0$. On the other hand, convicted criminals may have lower life-time wage earnings due to lost work-time (as a share of an individual's active work life) spent in prison, and/or by way of paying a fine (which may be increasing in the individual's potential wage rate) and/or by way of the labor market's negative wage response to workers who are ex-convicts, suggesting $\sigma_0<1$.

The μ-parameters represent returns from crime that are "uncorrelated" with the individual's potential wage. For certain crimes, such as tax evasion, the "wage-sensitivity" parameters σ would typically be positive while the μ-parameters may be close to zero, while for other crimes, such as full-time drug dealing, the σ-parameters

the cost of maintaining a given probability π is an increasing function of x, that is, we then allow for endogeneous law enforcement capacity.

would typically be close to zero and the μ-parameters positive. For $\mu_0 < 0$, $-\mu_0$ may be thought of as a fixed fine to be paid by convicted individuals. Unlike μ_1 and σ_1, μ_0 and σ_0 can be more or less directly influenced by policy (the length of sentence and size of penalty).

So far, we have focused on purely economic consequences of work and crime. We now turn to attitudes towards criminal activity. We assume that there is a social norm against engaging in criminal activity and that the intensity of this norm is endogenous. Individuals who choose crime loose utility from adherence to this social norm, and this utility is non-increasing in the expected crime rate. In other words, when many others are expected to adhere to the norm, a deviation from the norm results in a (weakly) larger utility loss that when only few others are expected to adhere to the norm.

Moreover, workers and offenders alike receive a disutility from criminal activity in society at large. This disutility thus represents the externality that certain criminal activities give rise to. An example is the fear, violence, restrictions on personal freedom and demoralizing influence on the young that illegal trade in drugs or weapons usually give rise to.

In sum, and more precisely, the expected utility associated with each of the two choices, *work* and *crime*, respectively, are

$$U_W = u\left(w - \tau w, 1\right) + av\left(x^e\right) - \psi\left(x^e\right) \tag{2}$$

and

$$U_C = \left(1 - \pi\right) u\left(\sigma_1 w + \mu_1, h_1\right) + \pi u\left(\sigma_0 w + \mu_0, h_0\right) - \psi\left(x^e\right), \tag{3}$$

where $\pi \in [0, 1]$ is the probability of being caught and convicted, which we take to be the same for all criminals. We assume the *consumption utility function* $u : \mathbb{R}^2_{++} \to \mathbb{R}$ to be continuous and strictly increasing in both arguments, the first being consumption of goods, the second, h, consumption of *leisure* (or, more generally, any relevant "material" attribute that distinguishes the three different "roles" each individual may be in: as a worker, a criminal at large, or as a caught and convicted criminal). We normalize leisure of a non-criminal worker to one unit. Moreover, we assume the *norm-adherence utility function* $v : [0, 1] \to \mathbb{R}_+$ to be strictly positive, continuous and non-increasing, and the *utility weight* $a \in \mathbb{R}$ that the individual places on norm adherence to be positive or negative, and the *externality function* $\psi : [0, 1] \to \mathbb{R}_+$ to be nonnegative, continuous and non-decreasing. For the sake of analytical tractability, we assume all individuals to have the same subutility functions u, v and ψ, while they may differ with respect to their individual potential wage, w, and degree of norm attachment, a. Moreover, we assume that a is a positive affine function of the punishment probability π:

$$a = \alpha + \beta\pi, \tag{4}$$

for $\alpha \in \mathbb{R}$ and $\beta \geq 0$. This representation can be derived from a psychological consideration, namely, the distinction between *guilt* and *shame*, that is, whether or not the disutility from norm deviation depends on own or others' disapproval. The parameter α represents the degree to which norm adherence is driven by guilt—as in the case of an internalized norm, where one's disutility from deviating is insensitive to whether or not others observe the deviation. By contrast, the parameter β represents the degree to which norm adherence is driven by *shame*—as in the case of an externally sanctioned norm, where one's disutility from deviating is sensitive to whether or not others observe the deviation.

To see that this distinction indeed gives rise to (4), temporarily replace the current assumption of a norm-adherence utility *gain*, $av(x)$, when choosing "work", with an equally large (and thus behaviorally equivalent) norm-adherence utility *loss*, from choosing "crime". To a criminal who is not caught and convicted, let this loss be $\alpha v(x)$, while the loss to a criminal who is caught and convicted is $(\alpha + \beta) v(x)$. Assuming that only the criminal knows about his or her crime when not caught and convicted, the utility loss in the first case represents guilt while that in the second case represents both guilt and shame—assuming that others in society learn about the crime of a convicted criminal. This alternative utility representation results in exactly the same utility *difference* between "work" and "crime" as in the present model, under equation (4), and is hence behaviorally equivalent with the present formulation.

Finally, note that the expected crime rate, x^e, is the same in equations (2) and (3). It is thus independent of the individual's own choice. This simplifying assumption is reasonable in a large population, individuals then arguably neglect the effect of their own choice on the crime rate in society at large. More exactly, viewing the utility from norm adherence as a function of the expected population share of *others* (in society at large, or in one's peer group) who choose crime, but the externality as a function of the total population share of criminals, the argument x^e in the term $av(x^e)$ should be interpreted as the population share of others who choose *crime*. If the total population, which we treat as a continuum, instead were finite, say of size N, then we should have $\psi(x^e + 1/N)$ instead of $\psi(x^e)$ in equation (3). However, by continuity of ψ, the difference between these two function values tend to zero as

$N \to 0$. Thus, our analysis also applies approximately to finite but large populations.

We will treat the quadruple $(\pi, \sigma_0, \mu_0, \tau)$ as a vector of policy instruments, perceived as fixed and given by individuals when they decide whether or not to engage in criminal activity.[2]

3. Equilibrium crime rates

If each individual chooses the alternative with the highest expected utility, then the choice *crime* is optimal for an individual with norm attachment a and potential wage w if and only if

$$u(w - \tau w, 1) + av(x^e) \leq (1 - \pi) u(\sigma_1 w + \mu_1, h_1) + \pi u(\sigma_0 w + \mu_0, h_0). \qquad (5)$$

For an individual with potential wage w who expects a crime rate x^e there thus exists a unique critical degree of norm attachment, $a^o(w, x^e)$, such that it is optimal for the individual to choose *crime* if and only if the individual's norm attachment a does not exceed this critical degree, which is determined from indifference in (5):

$$a^o(w, x^e) = \Delta u(w) / v(x^e), \qquad (6)$$

where

$$\Delta u(w) = (1 - \pi) u(\sigma_1 w + \mu_1, h_1) + \pi u(\sigma_0 w + \mu_0, h_0) - u(w - \tau w, 1). \qquad (7)$$

In words, $\Delta u(w)$ is the individual's expected *consumption utility gain* from *crime*, as compared with *work*, given the tax rate τ, punishment probability π, and punishment (σ_0, μ_0). In other words, $a^o(w, x^e)$ is the ratio between the expected consumption utility gain from crime—that depends on the individual's potential wage w as well as on the (here fixed) policy parameters τ and π—and the utility from norm adherence—that depends on the expected crime rate x^e.

We note that the critical degree of norm attachment, $a^o(w, x^e)$, is non-decreasing and continuous in the expected crime rate, x^e. Moreover, this critical degree of norm attachment is positive if and only if crime results in a higher expected consumption utility than work. We also note that the externality, not surprisingly, is irrelevant to individuals' decisions; it depends only on the expected population share x^e of criminals (which we assumed to be independent of the individual's own decision, see above).

---

[2] Arguably also h_0, the amount of leisure that a caught and convicted criminal has, is influenced by policy. However, we will not analyze this aspect.

We assume that all individuals simultaneously choose between work and crime, for a given tax rate τ and punishment probability π. By definition, a profile of such individual choices constitutes a Nash equilibrium if and only if every individual's choice is optimal, given the others' choices. A crime rate x will be called an equilibrium crime rate if it is consistent with some Nash equilibrium profile. Let Φ_π be the cumulative probability distribution function (c.d.f.) of "types" (a, w) in the population, where we recall that $a = \alpha + \beta\pi$, and let x be the associated crime rate, that is, the population share that chooses *crime*.[3] We then have:

Proposition 1. *A crime rate $x \in [0, 1]$ is a Nash equilibrium crime rate if and only if it satisfies the fixed-point equation (8), where $F : [0, 1] \to [0, 1]$ is defined in equation (9). There exists at least one Nash equilibrium.*

$$x = F(x) \tag{8}$$

$$F(x) = \int_{w=0}^{+\infty} \int_{a=-\infty}^{a^o(w,x)} d\Phi(a, w) \tag{9}$$

Proof: For any expected crime rate $x^e \in [0, 1]$, the subset $C \subset T = \mathbb{R} \times \mathbb{R}_+$ of individual types (a, w) for whom *crime* is optimal is Borel measurable, by (6), and their population share equals $x = F(x^e)$. Hence, a strategy profile is a Nash equilibrium if and only if $F(x^e) = x^e$. Being continuous, the function F has at least one fixed point (by the intermediate value theorem applied to $f(x) = x - F(x)$, a continuous function with $f(0) \leq 0$ and $f(1) \geq 1$. **End of proof**.

In words, $F(x^e)$ is the crime rate that results if all individuals expect the crime rate to be x^e. The fixed-point equation (8) expresses the requirement that individuals' crime-rate expectations be fulfilled, $x^e = F(x^e)$, an assumption of "perfect foresight" or "rational expectations."

A solution x^* to equation (8) will be called an *equilibrium crime rate*. In the special cases when the subutility $v(x)$ from norm adherence is independent of the crime rate x, the right-hand side in (8) is a constant and the equilibrium crime rate *unique*. However, the function F is in general non-decreasing, and therefore multiple equilibrium crime rates is a possibility. Figure **??** is drawn for a logistic norm-adherence utility function v and exponentially distributed norm attachments

[3] For any expected crime rate x^e the subset of individuals who will choose *crime* is Borel measurable, by (6), and hence their population share x is well-defined. The function F is continuous and hence has at least one fixed point.

a, under the further assumption that every individual has the same potential wage w (the numerical specifications of all diagrams are provided in the appendix at the end of the paper).

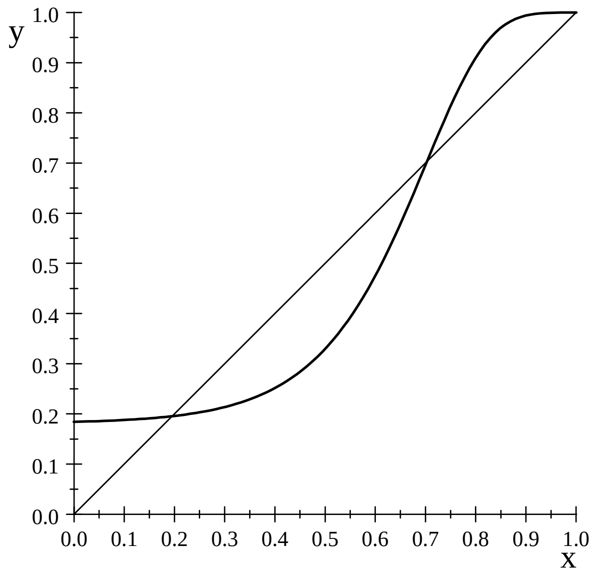

Figure 1: The equilibrium crime rate equation.

In this example, there are three equilibrium crime rates, $x \approx 0.2$, $x \approx 0.7$ and $x \approx 1$. The intuition for the multiplicity is that if the crime rate is low (high) then the utility from norm adherence — the guilt or shame from criminal activity — is strong (weak). Consequently, few (many) individuals choose to live off crime. Thus, two societies with the same potential wage, norm attachment distribution, penal code and police monitoring may differ significantly in their crime rates. That the highest equilibrium crime rate is close to 1 depends on $\Delta u(w)$ being positive, that is, that "crime pays" in terms of expected utility from income, and that the utility from norm attachment is close to zero when the crime rate is close to one (see Figure **??**). We also note that the intermediate equilibrium crime rate is unstable with respect to perturbations of expectations about others' behavior: if an individual expects the crime rate to be slightly higher (lower), then it is optimal to choose crime (work).

Equation (8) also defines the *Nash equilibrium correspondence* that maps parameter combinations to the associated (non-empty) set of equilibrium crime rates. We illustrate this correspondence by means of a diagram that we will discuss in two distinct cases: one when norm adherence is driven solely by guilt ($\beta = 0$), the other when norm adherence is driven solely by shame ($\alpha = 0$). Consider Figure 2 below.

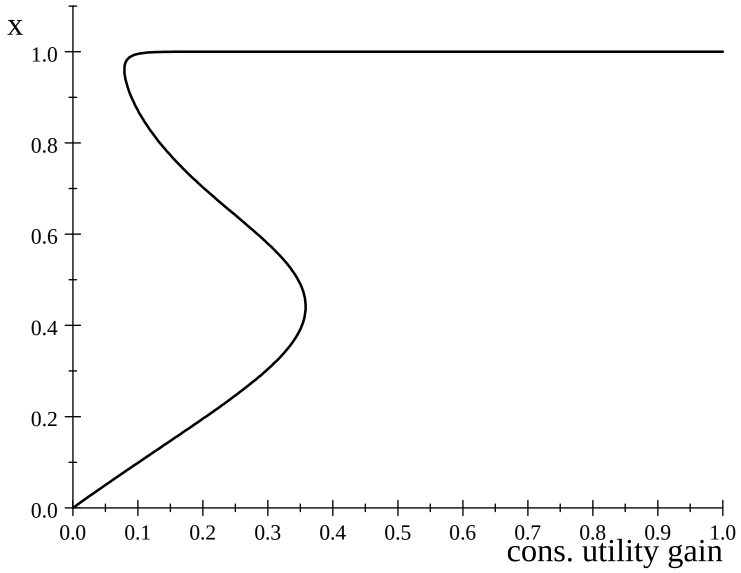

Figure 2: The equilibrium crime rate correspondence, mapping consumption utility gains to crime rates.

Imagine first that norm adherence is driven by guilt, that is, that $\beta = 0$. The diagram then illustrates the equilibrium correspondence that maps the expected consumption utility gain from crime, $\Delta u(w)$, defined in equation (7), to equilibrium crime rates, for α exponentially distributed. We see that the correspondence has a fold, and that it rises from zero to one as $\Delta u(w)$ increases from zero to 0.4. For values of $\Delta u(w)$ between approximately 0.1 and 0.37, there are three equilibrium crime rates, at each end-point of this interval there are two equilibria, and outside the interval a unique equilibrium. It is the endogeneity of the strength of the social norm, as represented by the function v, that causes the dramatic fold in the equilibrium correspondence. In the classical case without a social norm ($\alpha = \beta = 0$ for all individuals), the equilibrium correspondence is just a step function, jumping up from zero to one as the consumption utility gain from crime runs from negative to positive. Hence, besides potentially causing a fold — multiple equilibria — the social norm against crime, as modelled here, reduces the equilibrium crime rate when this utility gain is small and positive.

This diagram also provides some comparative statics insights. First, an increase in the punishment probability π decreases the expected consumption utility to crime, while an increase in the income tax rate τ increases it. Hence, as the population attaches less and less value to norm attachment, or as crime pays better and better, the equilibrium crime rate will increase, from almost 0% to almost 100%, but not continuously; there has to be at least one discrete upward jump as the consumption utility gain from crime $\Delta u(w)$ increases gradually from below 0.1 to above 0.37, approximately.

Let us make a heuristic thought experiment based on this diagram.[4] Suppose that the government gradually increases the punishment probability π, which decreases $\Delta u(w)$, *ceteris paribus*, starting from a value to the right of the fold, like $\Delta u(w) = 0.5$, and ending at a value to the left of the fold, such as $\Delta u(w) = 0.05$. The equilibrium crime rate will then decrease from about 100% to about 7%. According to the diagram, the crime rate necessarily makes a downward "jump" at some intermediate punishment probability. A gradual policy change thus results in a sudden and drastic fall in the value that people attach to the norm "not to engage in criminal activities" and thus also to their aggregate behavior. Reversing the thought experiment, that is, gradually reducing the punishment probability such that $\Delta u(w)$ is increased from $\Delta u(w) = 0.05$ to $\Delta u(w) = 0.5$, will likewise necessarily result in an upward jump at some point in time. If expectations-formation has inertia, in the sense that the equilibrium crime rate is expected to change only gradually when the punishment probability is changed marginally (whenever this is compatible with aggregate behavior), then we will have *hysteresis*: the upward jump will take place at a higher value of $\Delta u(w)$ (i.e. lower punishment probability) than the downward jump. In other words, if we think of the norm not to engage in crime as a form of social capital, then a gradual increase in the punishment probability leads to appreciation of that capital, while a gradual decrease in the punishment leads to depreciation of that capital, with hysteresis.

So far, we assumed that norm adherence was driven by pure guilt. Next, imagine that norm adherence is driven by pure shame, that is, $\alpha = 0$. Also then would the equilibrium correspondence have a graph like that in Figure **??**. However, now it would be the equilibrium correspondence that maps $\Delta u(w)/\pi$ (on the horizontal axis) to equilibrium crime rates (on the vertical axis).

[4]See Lindbeck, Nyberg and Weibull (1999, 2003) for discussions of social norm dynamics.

3.1. Crime earnings uncorrelated with potential wage. We here study in some more detail the case when the disposable income to criminals is independent of the potential wage rate. This may be the case of full-time illegal trade in drugs or weapons. More precisely, we now study cases when $\sigma_1 = \sigma_0 = 0$. For this particular study, we assume that all individuals have the same degree of norm attachment, $a = \alpha + \beta\pi$, but that wages w are distributed according to some cumulative distribution function G.

For any individual with norm attachment $a = \alpha + \beta\pi$ in a society with crime rate x there thus exists a unique critical wage value, $w^o(x)$, such that it is optimal to choose *crime* if and only if the individual's potential wage w does not exceed this critical value. The critical wage is determined from the indifference in the decision equation (5) which yields

$$w^o(x) = \frac{1}{1-\tau} u^{-1} \left[(1-\pi) u(\mu_1, h_1) + \pi u(\mu_0, h_0) - (\alpha + \beta\pi) v(x) \right], \qquad (10)$$

where u^{-1} is the inverse of the function $u(\cdot, 1) : \mathbb{R}_{++} \to \mathbb{R}$.

The equilibrium equation (8) can be written now as

$$x = G\left[\frac{1}{1-\tau} u^{-1} \left[(1-\pi) u(\mu_1, h_1) + \pi u(\mu_0, h_0) - (\alpha + \beta\pi) v(x) \right] \right]. \qquad (11)$$

This equation defines a correspondence that maps each policy triplet (π, μ_0, τ), where $\mu_0 > 0$ is a caught and convicted criminal's consumption, to the associated set of equilibrium crime rates x (recall that we fixed σ_0 at zero).

In order to facilitate the analysis of this correspondence, we assume equal leisure in all three roles of an individual, logarithmic consumption utility function, and that the wage density $g(w) = G'(w)$ is positive for all positive wages $w > 0$. The c.d.f. G then has an inverse, and equation (11) can be re-written as

$$G^{-1}(x) \exp\left[(\alpha + \beta\pi) v(x) \right] = \frac{1}{1-\tau} \exp\left[(1-\pi) \ln \mu_1 + \pi \ln \mu_0 \right]. \qquad (12)$$

The right-hand side of (12) is a continuous function of the tax rate τ, the consumption μ_0 of convicted criminals, and the punishment probability π, strictly increasing in the first **two** and strictly decreasing in the third. Likewise, the left-hand side is a continuous function of the crime rate x and the punishment probability π. It is non-decreasing in the latter (strictly increasing iff $\beta > 0$), and strictly increasing in x if the utility from norm-adherence is relatively insensitive to the crime rate. In such cases, the equilibrium crime rate x is uniquely determined by τ, μ_0 and π. Moreover, the equilibrium crime rate is then increasing in the tax rate — since a higher income tax rate makes work less attractive — and in the consumption level of convicted criminals, but decreasing in the punishment probability, for obvious reasons. More precisely, by way of differentiation of the left-hand side in (12) with respect to x, one obtains:

Proposition 2. *The equilibrium crime rate is unique under condition [M] below. It is increasing in τ and μ_0, and decreasing in π.*

$$[M] \qquad v'(x)\, g\left(G^{-1}(x)\right) G^{-1}(x) > -\frac{1}{\alpha + \beta\pi} \qquad \forall x$$

By contrast, if the left-hand side in equation (12) is not increasing in x, then certain policies (π, μ_0, τ) admit multiple equilibrium crime rates x. See Figure 3, drawn for a case when the norm-adherence utility function is logistic in the crime rate and the potential wage w is distributed according to a Weibull $(0, \frac{1}{2}, 2)$ distribution. The wavy curve is the graph of the left-hand side in equation (12), viewed as a function of x, and the straight line gives the value of the right-hand side of the same equation.

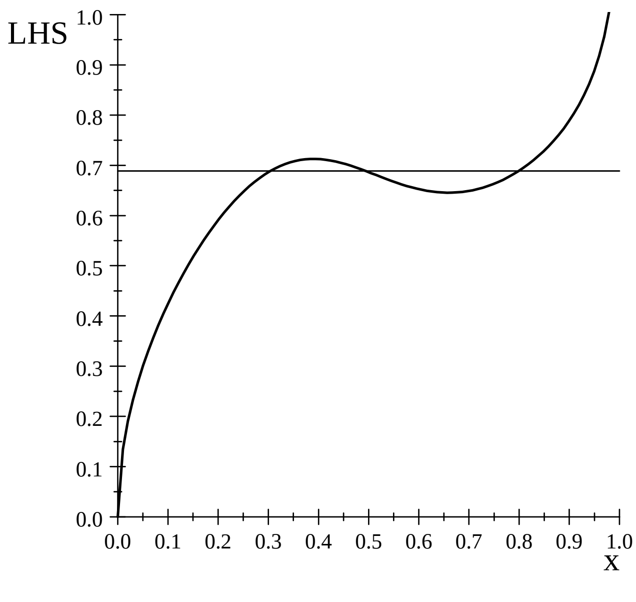

Figure 3: The left-hand side in equation (12) as a function of the crime rate x.

Figure 4 below shows the equilibrium correspondence that maps punishment probabilities π to equilibrium crime rates x when the income tax rate τ and consumption μ_0 of convicted criminals are kept fixed (at the levels $\tau = 0.25$ and $\mu_0 = 0.1$) as in the specification of Figure 3). The thicker and folded curve is the graph of the equilibrium correspondence when all individuals have norm attachment parameters $\alpha = 1$ (guilt) and $\beta = 1$ (shame), while the thinner curve corresponds to the case of zero norm attachment — the standard economics model of crime and punishment (initiated by Becker (1968)). We see that the presence of a social norm against committing crimes suppresses the equilibrium crime rate, less at low punishment probabilities and more at medium to high punishment probabilities, and that the endogeneity of the social norm causes again a dramatic fold in the equilibrium correspondence. In particular, condition [M] is violated in this example, since, by proposition 2 the condition is sufficient for the non-existence of folds in this correspondence.

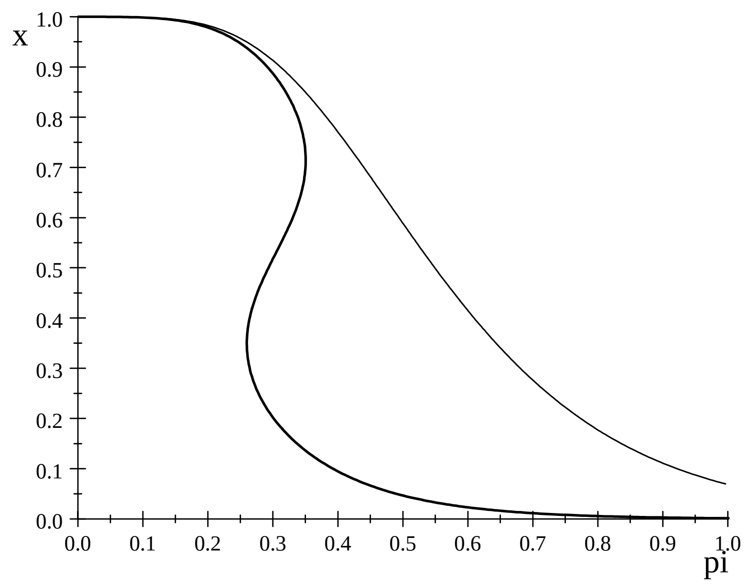

Figure 4: The equilibrium correspondence that maps punishment probabilites to crime rates.

Let us make again a heuristic thought experiment based on this diagram. Suppose that the government gradually increases the punishment probability π, starting from a value to the left of the fold, such as $\pi = 0.25$, and ending at a value $\pi = 0.5$ to the right of the fold. The crime rate will then decrease from about 95% to about 10%. According to the diagram, the crime rate necessarily makes a downward "jump" at some intermediate punishment probability. A gradual policy change thus results in a sudden and drastic fall in the value that people attach to the norm "not to engage in drug dealings" and thus also to their aggregate behavior. If expectations-formation shows inertia then we will have again *hysteresis*: the upward jump will take place at a lower punishment probability than the downward jump.[5]

3.2. Crime earnings correlated with potential wage. We here study the case when criminals' disposable income is positively related to their potential wages, as may be the case with tax evasion. Indeed, Allingham and Sandmo (1972) emphasize that tax evasion may be subject to shame or stigma. Tax evaders may be quite embarrassed when they are caught under-reporting their income. We here outline how our model can be applied to such forms of illegal behavior.

Individuals face a binary choice when reporting their income to the tax authorities: to either report the full income, w, or only the fraction ρw, where $\rho \in (0,1)$ is fixed and given. Reported incomes are taxed at a constant rate $\tau \in (0,1)$. Thus, an individual with income w who reports truthfully pays τw in taxes while an under-reporting individual pays only $\tau \rho w$. If an individual is caught under-reporting, then she must pay the amount withheld, $(1-\rho)\tau w$, and a fraction $\gamma \in (0,1)$ of her remaining disposable income (or, equivalently, spend time in prison and thereby lose work time). We will call γ the *penalty rate*. Hence, in the notation of section 2:

$$\sigma_1 = 1 - \rho\tau, \;\; \sigma_0 = (1-\gamma)(1-\tau) \;\text{ and }\; \mu_1 = \mu_0 = 0.$$

We note that $\sigma_0 < 1 - \tau < \sigma_1$; the disposable income from under-reporting and being caught and convicted is lower than the disposable income from reporting truthfully, which in its turn is lower than the disposable income from under-reporting and not being caught and convicted.[6] The crime rate $x \in [0,1]$ is now the population fraction of tax evaders. We assume leisure to be the same in all three roles of an individual, $h_0 = h_1 = 1$, and suppress this constant argument in the consumption utility function.

The expected utility associated with each of the two choices, *report truthfully* and

[5] The two probabilities being approximately $\pi = 0.26$ and $\pi = 0.35$

[6] Parameters are assumed to be such that $\sigma_0 > 0$, that is, caught and convicted tax evaders are left with a positive disposable (life time) income.

underreport, respectively, are thus

$$U_W = u\left[(1-\tau)\,w\right] + av\,(x) - \psi\,(x) \tag{13}$$

and

$$U_C = (1-\pi)\,u\left[(1-\rho\tau)\,w\right] + \pi u\left[(1-\gamma)\,(1-\tau)\,w\right] - \psi\,(x) \tag{14}$$

where $a \in \mathbb{R}$ is the utility weight that the individual places on adherence to the social norm to report truthfully.

In the special case of logarithmic subutility from consumption, the potential wage cancels out when comparing U_W with U_C, and we obtain

$$\Delta u\,(w) \equiv (1-\pi)\ln\left(\frac{1-\rho\tau}{1-\tau}\right) + \pi\ln\,(1-\gamma)\,.$$

from (7). In other words, individuals' choices are independent of their potential wages, and the consumption utility gain from crime is increasing in the tax rate τ and decreasing in the punishment probability π.

We briefly consider this case. First, suppose that all individuals have the same positive norm attachment a. Then the crime rate is unique, $x^* = 0$, if $\Delta u\,(w) \leq av\,(1)$: even if all others were to defect from the norm, truthful reporting is preferable. Likewise, $x^* = 1$ is the unique equilibrium crime rate if $\Delta u\,(w) \geq av\,(0)$: even if no one else were to defect from the norm, under-reporting is preferable. However, in the remaining case, that is, when $av\,(1) < \Delta u\,(w) < av\,(0)$, there are three equilibria, namely, $x^* = 0$, $x^* = 1$, and $x^* = v^{-1}\left[\Delta u\,(w)\,/a\right]$. The reason for this multiplicity is simple. If the crime rate is expected to be $x^* = v^{-1}\left[\Delta u\,(w)\,/a\right]$, then all individuals are indifferent between reporting truthfully and under-reporting. Hence, this crime rate is an equilibrium. If the crime rate is expected to be lower (higher), then the unique optimal choice is truthful reporting (under-reporting).[7]

Secondly, suppose individuals differ in their norm attachment, and suppose this is driven solely by shame. Indeed, it seems reasonable to conjecture that shame is more prominent than guilt in the present case of tax evasion, since tax evasion convictions easily become public and may cause considerable social stigma. Hence, we now have $\alpha = 0$, while β is distributed according to some c.d.f. B. Equation (8) then becomes

$$x = B\left(\frac{1}{v\,(x)}\left[\frac{1-\pi}{\pi}\ln\left(\frac{1-\rho\tau}{1-\tau}\right) + \ln\,(1-\gamma)\right]\right), \tag{15}$$

[7] This also shows that the equilibrium crime rate $x^* = v^{-1}\left[\Delta u\,(w)\,/a\right]$ is unstable with respect to perturbations of expectations about others' behavior.

a fixed-point equation with a graph identical to that in Figure **??**, under suitable parameter specifications.[8]

The equilibrium equation (15) also defines a correspondence that maps policies, (γ, π, τ), to the associated set of equilibrium crime rates, x. Figure 5 below shows the the graph of this equilibrium correspondence, for a fixed tax rate and punishment probability—that is, the correspondence from γ to x, given τ and π (for the same numerical specification as in the preceding footnote). The thinner curve corresponds to the case of no attachment to the social norm, that is when $B(z) = 0$ for all $z < 0$ and $B(z) = 1$ for all $z \geq 0$. In this case, all individuals have the same preferences and thus make identical choices. Therefore, the equilibrium crime rate drops from 1 to 0 as the penalty rate γ increases from below to above a certain critical value γ^o (here approximately 0.35, see appendix).

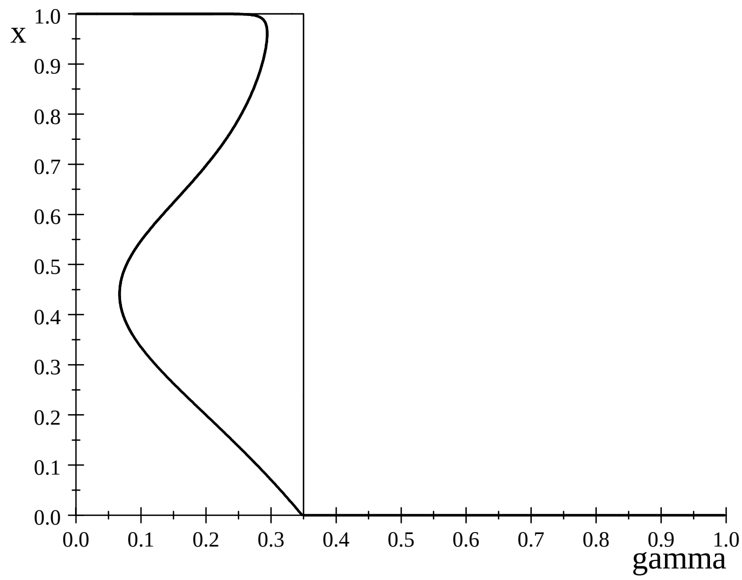

Figure 5: The equilibrium correspondence that maps penalty rates to equilibriium crime rates.

For all penalty rates $\gamma > \gamma^o$, the expected consumption utility from crime falls short of that from work, so no one then chooses tax evasion, even in the absence

[8]More precisely, if v is logistic, β exponentially distributed with mean value 1, $\gamma = 0.2$, $\rho = 0.2$, $\tau = 0.4$ and $\pi = 0.5$, then $\Delta u(w) = \frac{1-\pi}{\pi} \ln\left(\frac{1-\rho\tau}{1-\tau}\right) + \ln(1-\gamma) \simeq 0.2$, just as in Figure **??** (see appendix).

of a social norm against such behavior. Hence, the unique equilibrium tax evasion rate is zero, both with and without attachment to the social norm. For sufficiently low penalty rates (below 0.07, approximately), the expected consumption utility from crime is so high that all individuals choose crime, even with norm attachment. For intermediate penalty rates, $0.07 < \gamma < 0.3$, there are three equilibrium crime rates, in the case with a positive norm attachment. The intermediate equilibrium crime rate is unstable under expectation perturbations, and thus does not seem plausible as a prediction. The highest equilibrium crime rate would be close to one while the lowest equilibrium crime rate lies in the interval between 0 and 0.45, and decreases smoothly as the penalty rate increases.

Suppose that initially the penalty rate is high (above 0.35), and then gradually reduced, under the social norm. With inertia in expectations formation, we would then see how the tax evasion rate would gradually rise and then suddenly jump up, from about 45% to 100% , as the penalty rate gradually falls below 0.07. Compared with the case with no attachment to the social norm, this jump thus takes place at a much lower penalty rate. The social norm keeps the crime rate down as long as the economy is locked in on the lower equilibrium. Had the penalty rate thereafter been gradually increased, then hysteresis would be observed: the evasion rate would remain very high until the penalty had reached the value 0.3.

A society that is locked in at the lowest of the three equilibria thus has a lower rate of tax evasion than can be explained by economic incentives alone. Frey and Feld (7) point out that under purely economic incentives, as in the original Allingham-Sandmo model, tax evasion should be much higher than what is empirically observed, given the low level of deterrence in most countries.[9] The present model can potentially explain such empirical observations. The model provides an explanation of why tax evasion empirically seems to be less pervasive than can be explained by purely economic incentives.

4. Crime-deterrence policy as political equilibrium

The model can be used for policy analysis, in particular for analyses of alternative crime deterrence measures. This can be done normatively—to see what policies maximize a given welfare function—or positively—in order to predict policy as an outcome of a democratic voting process.

[9] Frey and Feld (7) argue that to achieve the high compliance rates of 80% observed (respectively low tax evasion rates) the Arrow-Pratt risk aversion measure would have to be unreasonably high (around 30) in the original Allingham-Sandmo model to match what is currently observed for the United States and Switzerland. See also Sandmo (13) and Pommerehne and Weck-Hannemann (11).

As noted above, the present model contains three potential instruments for crime deterrence: the punishment probability π, that is, the probability that a given criminal be caught and convicted, and, indirectly σ_0 and μ_0, where we recall that $\sigma_0 w + \mu_0$, is the income to an individual with potential wage rate w who has been caught and convicted. Hence, σ_0 in part depends on the extent to which criminals also work and pay taxes, but also in part on the duration of imprisonment (in terms of lost lifetime wage earnings), and $-\mu_0$ may be interpreted as a fixed fine to be paid by an individual who is caught and convicted (if the fine in part depends on the individual's potential wage rate, then this will also influence σ_0).

We illustrate these possible uses of the model by way of discussing an example. But first some general points.

4.1. Public budget balance. Let $C(\pi, \sigma_0, \mu_0, x)$ be the cost to government of keeping the punishment probability at the level π and the duration of imprisonment and the fines such that they result in σ_0 and μ_0, when the crime rate is x. Plausibly, this cost is increasing in all arguments.

Suppose that the income tax τ is the sole source of revenue for the government. Given the government's commitments to other public expenditures, $E \geq 0$, the government budget is then balanced if and only if

$$R(\pi, \sigma_0, \mu_0, x) = C(\pi, \sigma_0, \mu_0, x) + E, \tag{16}$$

where the left-hand side, $R(\pi, \sigma_0, \mu_0, x)$, is the total tax revenue collected from workers (those who choose *work* and those who choose *crime* but also work part time), when the crime rate is x. While tax revenues are plausibly decreasing and continuous in the crime rate, at fixed policy parameters, expenditures are increasing and continuous. Therefore, if other public spending, E, is not excessive in relation to the income tax rate, there will exist a unique crime rate that balances the government budget (where the downward-sloping revenue curve intersects the upward-sloping expenditure curve). A quintuple $s = (\pi, \sigma_0, \mu_0, \tau, x)$ that satisfies the budget equations (16) and the equilibrium equation (8) will be called a *balanced equilibrium state* of the economy, and we denote this set S^*.

A normative approach to crime deterrence policy can thus be developed by way of maximization of a welfare function over the set of balanced equilibrium states, defined by these two equations.

4.2. Political equilibrium. By a *political equilibrium* we mean a balanced equilibrium state such that no other balanced equilibrium state is preferred by a majority of voters. More exactly, the voting situation for each individual may be thought of as a vote between a current policy $p = (\pi, \sigma_0, \mu_0, \tau)$, and some "opposing" policy $p' = (\pi', \sigma'_0, \mu'_0, \tau')$, where the current crime rate x is uniquely determined by the condition that $s = (\pi, \sigma_0, \mu_0, \tau, x)$ be a balanced equilibrium state, and where $(\pi', \sigma'_0, \mu'_0, \tau')$ is such that $s' = (\pi', \sigma'_0, \mu'_0, \tau', x')$ is a balanced equilibrium state (where again x' is uniquely determined by p'). All voters expect the current crime rate x under the incumbent policy p, and they all expect some crime rate x' under the alternative policy p'.

The expected utility to an individual with potential wage rate $w > 0$ and norm attachment parameters α and β in any state $s = (\pi, \sigma_0, \mu_0, \tau, x)$ is

$$U(s, w, \alpha, \beta) = \max\{U_W, U_C\},$$

where

$$U_W = u(w - \tau w, 1) + (\alpha + \beta\pi) v(x) - \psi(x) \tag{17}$$

and

$$U_C = (1 - \pi) u(\sigma_1 w + \mu_1, h_1) + \pi u(\sigma_0 w + \mu_0, h_0) - \psi(x). \tag{18}$$

In other words, each individual anticipates to make an optimal individual choice were the state s to materialize.

We define a policy $p = (\pi, \sigma_0, \mu_0, \tau)$ to be a *political equilibrium policy*, or an *unbeatable* policy under majority rule, if the "current" state $s = (\pi, \sigma_0, \mu_0, \tau, x)$ is a balanced equilibrium state and is preferred by a (weak) majority over any alternative balanced equilibrium state $s' = (\pi', \sigma'_0, \mu'_0, \tau', x')$, that is, if

$$\int_{w=0}^{+\infty} \int_{a=-\infty}^{+\infty} H[U(s, w, \alpha, \beta) - U(s', w, \alpha, \beta)] \, d\Phi(a, w) \geq 1/2 \quad \forall s' \in S^*, \tag{19}$$

where H is the indicator function defined by $H(x) = 1$ for $x \geq 0$ and $H(x) = 0$ for $x < 0$.[10]

A positive approach to crime deterrence policy can thus be developed by way of identifying the (potentially empty) subset $S^0 \subset S^*$ of balanced equilibrium states that satisfies (19). Given this subset, the associated subset P^0 of political equilibrium policies is defined as those policies $(\pi, \sigma_0, \mu_0, \tau)$ for which $s = (\pi, \sigma_0, \mu_0, \tau, x) \in S^0$.

We illustrate this abstract machinery by way of an example.

[10] In this definition, the preference may be required to be strict or weak and likewise with the majority. We here take both to be weak.

4.3. Tax-evasion: an example. Suppose that all individuals have logarithmic consumption utility and that their norm attachment is driven solely by shame ($\alpha = 0$). Suppose, moreover, that the shame parameter β is distributed according to some continuous cumulative probability distribution function B. We focus on the special case of tax evasion when $\rho = 0$, that is, when all tax evaders report zero income. Moreover, assume that the penalty γ paid by convicted tax evaders is not a source of revenues to the government (more realistically, part of these penalties would constitute government income). The punishment probability π is assumed to be fixed, and, at that punishment probability, the cost of maintaining and enforcing the legal code is assumed to be an affine increasing function of the crime rate: $C = C_0 + cx$, where $C_0 > 0$ is the fixed cost and $c > 0$ the constant marginal cost. This is the case if the capacity of the law enforcement system can be adjusted, at a cost, so as to keep the punishment probability at a given level.

Hence, for any crime rate x, government revenues emanate from only one source, namely from those who work and pay taxes, and these revenues have to finance its expenses:

$$(1 - x)\,\tau\bar{w} = C_0 + cx + E, \tag{20}$$

where $C_0 + E > 0$ is fixed, and $\bar{w}$ is the average wage. (Recall that individual decisions are independent of one's own wage.) The only remaining policy variables are τ, the income tax rate, and γ, the penalty rate paid by tax evaders. More precisely, a tax evaded with wage w pays his or her tax debt, τw, and from the remaining income, $(1 - \tau)\,w$, the share γ.

Combining the public budget balance equation with the equilibrium equation (15), we obtain the following equation in τ and γ:

$$\frac{\tau\bar{w} - C_0 - E}{\tau\bar{w} + c} = B\left(\left[v\left(\frac{\tau\bar{w} - C_0 - E}{\tau\bar{w} + c}\right)\right]^{-1}\left[\ln(1-\gamma) - \frac{1-\pi}{\pi}\ln(1-\tau)\right]\right). \tag{21}$$

Note that if the penalty rate is maximal, $\gamma = 1$, then the consumption utility of a caught and convicted criminal is minus infinity. Hence, the crime rate is then zero and, by (20), the tax rate is

$$\tau_0 = \frac{C_0 + E}{\bar{w}}, \tag{22}$$

and where we recall that $\bar{w}$ is national income.[11] The policy $(\tau, \gamma) = (\tau_0, 1)$ is ideal for every worker, since the tax rate is minimal and the crime rate is zero, hence

[11] To see how this follows from equations (20) and (21), note that for $\gamma = 1$, the right-hand side of the latter equation becomes zero, and hence $\tau\bar{w} = C_0 + E$. Inserting this expression for $C_0 + E$ in the first equation, we obtain $-x\tau\bar{w} = cx$, which has $x = 0$ as its unique solution.

minimizing the cost of law enforcement and also maximizing the utility to a worker from norm adherence.

Under what conditions is the policy $(\tau, \gamma) = (\tau_0, 1)$ a political equilibrium? The answer is there should exist no alternative balanced-budget equilibrium policy under which a majority would be tax evaders and obtain a higher expected utility than when working and paying income taxes according to the tax rate τ_0. More precisely, let $\tilde{\beta}$ be the median value of the norm-attachment parameter β, and assume $\psi = 0$, that is, that there is no externality from tax evasion. Then

Proposition 3. $(\tau_0, 1)$ *is not a political equilibrium policy if and only if (23) and (24) hold for some policy* (τ, γ) *that satisfies (21).*

$$(1-\tau)(1-\gamma) \geq (1-\tau_0)^{1/\pi} e^{\tilde{\beta} v(0)} \tag{23}$$

$$\tau \geq 2\tau_0 + c/\bar{w} \tag{24}$$

Proof: The policy $(\tau_0, 1)$ cannot be beaten under majority vote by a policy that results in a majority of workers, since these would be better off under $(\tau_0, 1)$. Hence, only a policy (τ, γ) satisfying (21) and such that equation (20) gives $x \geq 1/2$ can beat $(\tau_0, 1)$. Note that $x \geq 1/2$ holds in a budget-balanced equilibrium if and only if (24) holds. (The latter inequality is obtained by solving for x in (20) and requiring $x \geq 1/2$.) Hence, a competing policy (τ, γ) results in $x \geq 1/2$ iff (24) holds, and such a policy beats $(\tau_0, 1)$ if and only if the expected utility to a tax evader is higher than when working and paying income tax τ_0 under policy $(\tau_0, 1)$. Using (17) and (18), the latter condition boils down to

$$(1-\pi)\ln w + \pi \ln((1-\gamma)(1-\tau) w) > \ln((1-\tau_0) w) + \beta\pi v(0)$$

or, equivalently,

$$\pi \ln((1-\gamma)(1-\tau)) > \ln(1-\tau_0) + \beta\pi v(0).$$

This inequality needs to hold for all $\beta \leq \tilde{\beta}$ in order for x to be at least $1/2$. This is condition (23). **End of proof.**

The proposition has certain comparative statics implications, which can be summarized as follows. Let

$$P_0 = \left\{(\tau, \gamma) \in [0,1]^2 : \text{eqs. (23-24) hold}\right\}.$$

We say that $(\tau_0, 1)$ is *more easily* a political equilibrium the smaller P_0 is. It follows from the above proposition that the policy $(\tau_0, 1)$ is more easily a political equilibrium the stronger the social norm is against tax evasion, the higher the probability is for a tax evader to be caught and convicted, and the higher is the marginal cost (as a share of national income) of law enforcement:

Corollary 1. *$(\tau_0, 1)$ is more easily a political equilibrium, the higher $\tilde{\beta} v(0)$ is, the higher π is, and the higher c/w is.*

Can any other policy be a political equilibrium? Since $(\tau_0, 1)$ is the ideal policy for workers, political equilibria with a majority of tax evaders are the only possible alternatives. In order to address this question, first note that (21) defines the penalty rate γ as a function of the tax rate τ. More exactly, equation (21) is satisfied if and only if $\gamma = g(\tau)$, where

$$g(\tau) = \max\left\{0, 1 - (1-\tau)^{\frac{1-\pi}{\pi}} \exp\left[B^{-1}\left(\frac{\tau - \tau_0}{\tau + c/\bar{w}}\right) v\left(\frac{\tau - \tau_0}{\tau + c/\bar{w}}\right)\right]\right\}. \tag{25}$$

Let $\hat{\tau}$ be the lowest tax rate that is compatible with a (weak) majority of tax evaders. By (20), $x \geq 1/2$ requires $\tau \geq \hat{\tau}$, where

$$\hat{\tau} = 2\tau_0 + c/\bar{w}. \tag{26}$$

Hence, a necessary condition for the existence of political equilibrium with a majority of tax evaders is $\hat{\tau} < 1$. Let

$$\tau^* = \min \arg \max_{\tau \in [\hat{\tau}, 1]} (1-\tau)^{\frac{1}{\pi}} \exp\left[B^{-1}\left(\frac{\tau - \tau_0}{\tau + c/\bar{w}}\right) v\left(\frac{\tau - \tau_0}{\tau + c/\bar{w}}\right)\right]. \tag{27}$$

This is the minimal tax rate among those that maximize tax evaders' expected utility in balanced-budget equilibrium where these constitute a weak majority. For continuous functions B^{-1} and v, the maximand is continuous and hence the set of maximzers is non-empty and compact (by Weierstrass' Maximum Theorem), so τ^* is then well-defined, granted $\hat{\tau} < 1$. Let $\gamma^* = g(\tau^*)$. This is the unique penalty rate that makes (τ^*, γ^*) a balanced equilibrium policy. If there is a political equilibrium with a majority of tax evaders, then the tax rate in that equilibrium needs to be τ^* and the penalty needs to be γ^* since otherwise policy (τ^*, γ^*) would defeat that policy under majority rule. Also the converse holds:

Proposition 4. *Suppose that* $\hat{\tau} < 1$ *and* B^{-1} *is continuous. Then* (τ^*, γ^*) *is a political equilibrium if and only if*

$$(1-\tau^*)(1-\gamma^*) \geq (1-\tau_0)^{1/\pi} e^{\tilde{\beta} v(0)}. \tag{28}$$

Proof: Suppose that $\hat{\tau} < 1$, B^{-1} is continuous and (28) holds. Then (τ^*, γ^*) results in a balanced equilibrium state with a (weak) majority of tax evaders, and, by (28) their expected utility is at least as high as under the optimal policy for workers. Hence, no policy (τ, γ) with $\tau < \hat{\tau}$ can obtain a majority against (τ^*, γ^*). But nor can any policy with $\tau > \hat{\tau}$ since τ^* maximizes tax evaders' expected utility across all balanced-budget equilibria where these constitute a weak majority. **End of proof.**

It is harder to discuss the comparative statics of this proposition than of the previous, since the candidate tax rate, τ^*, is more indirectly defined in terms of the primitives than is τ_0. However, in some special cases, the conditions of the proposition are fairly transparent.

We here focus on the special case when a majority attaches no weight to the social norm, that is, when $\tilde{\beta} = 0$, while the remaining minority attaches some positive weight $\beta_1 > 0$ on the norm. Let thus $x_1 = B(0) > 1/2$. The minority of individuals who care about the norm would be willing to report their incomes honestly even if the others don't and even if there were no penalty for tax evaders ($\gamma = 0$), granted that

$$\beta_1 v(x_1) \geq \left(1 - \frac{1}{\pi}\right) \ln(1-\tau), \tag{29}$$

where τ is the going tax rate. Suppose this is chosen such that the tax revenue from the "honest" minority is sufficient to finance all public expenditures, that is, $\tau = \tau_1$, where

$$\tau_1 = \frac{C_0 + cx_1 + E}{(1-x_1)\bar{w}} < 1.$$

In the absence of a penalty for tax evaders, the "dishonest" majority will choose tax evasion, and thus $(\tau_1, 0)$ is a political equilibrium. We see in (29) that, not surprisingly, this is easier to obtain the larger is the utility $\beta_1 v(x_1)$ that the honest derive from adhering to the norm of truthful income declaration when the population fraction $x_1 = B(0)$ violate the norm. The larger that utility is, the more vulnerable are the honest to exploitation by the dishonest.

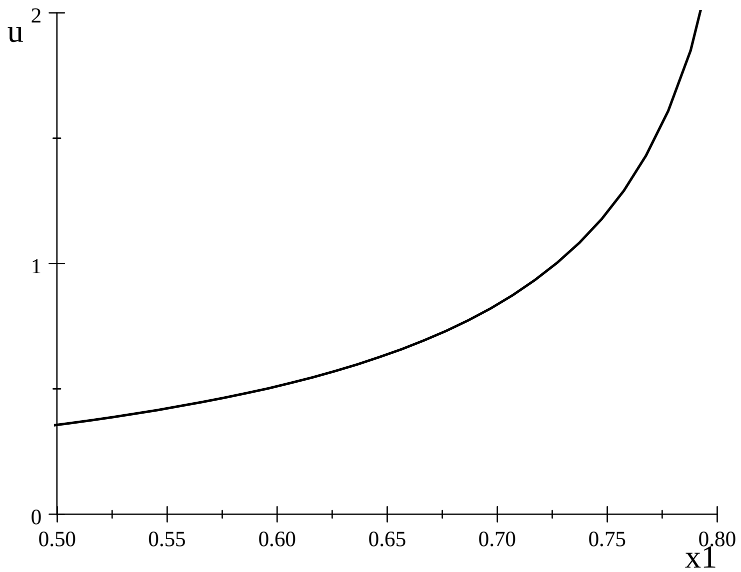

Figure 6: The political equilibrium correspondence that maps the share of tax evaders to the norm attachment of the honest.

We end by illustrating this graphically. Figure 6 plots the minimal utility value of norm adherence, $u = \beta_1 v(x)$, needed for a given population share $x = B(0) \geq 1/2$ of individuals with no norm attachment for the policy $(\tau_1, 0)$ to be a political equilibrium, see Appendix for the numerical specification. The higher the utility of norm adherence for the honest is, the larger is the share of tax evaders that can exploit them in political equilibrium.

5. Appendix

We here provide the numerical specifications used in the diagrams.

5.1. Figures 1 and 2. Suppose that all individuals have the same potential wage w, that their norm attachments, a, are exponentially distributed with mean $\mathbb{E}(a) = 1$, and that the norm-adherence utility function v is logistic. Figure A1 shows the graph of the logistic norm-adherence utility function

$$v(x) = \frac{1}{1 + \exp(8x - 4)}. \tag{30}$$

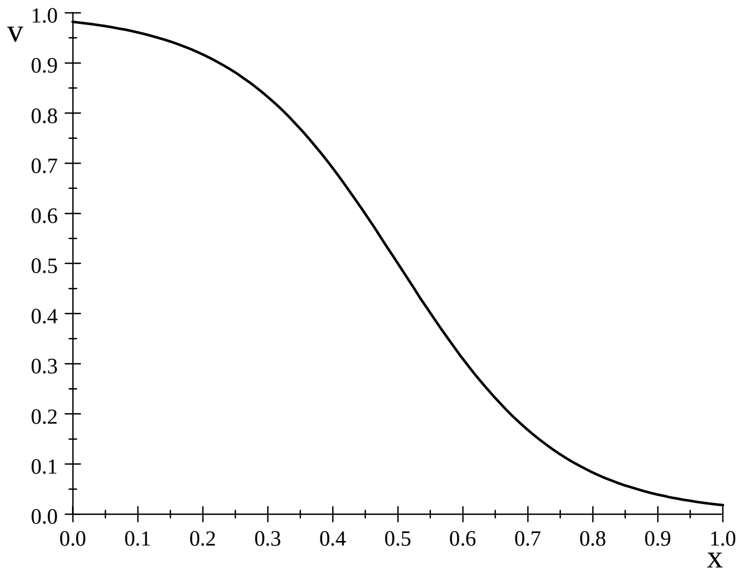

Figure A1

With this norm-adherence utility function and exponentially distributed norm attachment, the fixed-point equation (8) takes the (doubly exponential) form[12]

$$x = 1 - \exp\left(-\frac{\Delta u\left(w\right)}{\mathbb{E}\left(a\right)}\left[1 + \exp\left(8x - 4\right)\right]\right), \tag{31}$$

where $\Delta u\left(w\right)$ is given by (7). Figure 1 is drawn for $\Delta u\left(w\right) = 0.2$. Figure 2 uses (31) allowing $\Delta u\left(w\right)$ to vary implicitly with the crime rate x.

5.2. Figure 3 and 4. A Weibull (α, β, γ) c.d.f. $G\left(x\right)$ is zero for all $x < \alpha$, while for $x \geq \alpha$:

$$G\left(x\right) = 1 - \exp\left(-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right).$$

The inverse of this c.d.f., when $\alpha = 0$, is

$$G^{-1}\left(x\right) = \beta\left(\ln\left(\frac{1}{1-x}\right)\right)^{\frac{1}{\gamma}}.$$

Let the potential wage w be Weibull $(0, \frac{1}{2}, 2)$ distributed, see its density in Figure A2, and that the norm-adherence utility function v is logistic.

[12] The right-hand side is formally equivalent with a Gumbel, or doubly exponential, probability distribution function.

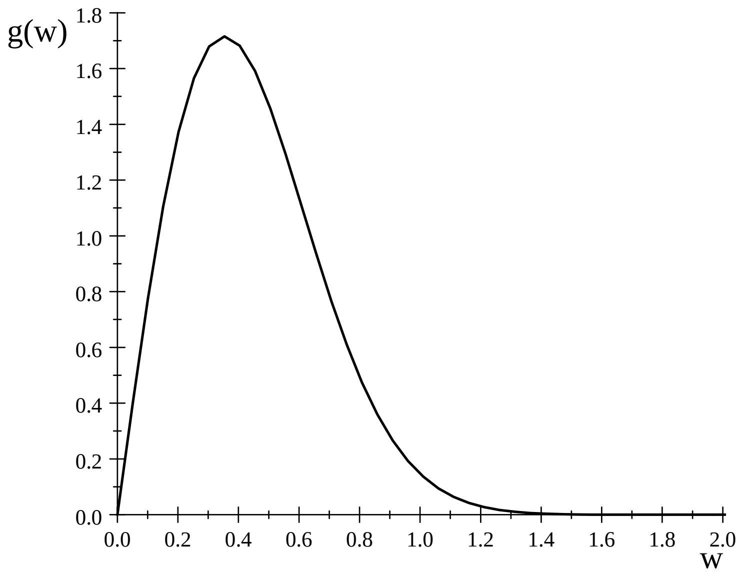

Figure A2

The wavy curve of Figure 3 is drawn for w distributed Weibull $(0, \frac{1}{2}, 2)$, $\alpha = 0.65$, $\beta = 1$ and $\pi = 0.35$. The other parameters were set such that $\tau = 0.25$, $\ln \mu_1 = 0.22$ (which corresponds to the consumption utility for a worker earning a potential wage in the upper tail of the wage distribution of 1.25 after tax) and $\ln \mu_0 = -2.3$ (the approximate consumption utility for a worker earning a potential wage in the lower tail of the wage distribution of 0.1 after tax). We have therefore that the straight line in Figure 3 is given by

$$\frac{1}{1-\tau} \exp\left[(1-\pi)\ln \mu_1 + \pi \ln \mu_0\right] = \frac{1}{1-0.25} \exp\left[0.65 \ln(1.25) + 0.35 \ln(0.1)\right] \simeq 0.69.$$

Figure 4 is based on this same specification maintaining fixed the tax rate $\tau = 0.25$ while letting the punishment probability π vary implicitly with x such that (8) is satisfied.

5.3. Figure 5. The thick curve of this diagram represents the case with a social norm that takes the form of pure shame ($\alpha = 0$ and β distributed exponentially with mean $\mathbb{E}(\beta) = 1$) and uses the following set up: v is a logistic norm-adherence utility function as in (30), $\gamma = 0.3$, $\rho = 0.2$ and $\tau = 0.25$. The fixed-point equation (**??**) is then

$$x = 1 - \exp\left(-\left(1 + \exp(8x - 4)\right)\left[0.23\left(\frac{1-\pi}{\pi}\right) - 0.35\right]\right)$$

which is the equation that implicitly relates x to π and generates the thick curve in Figure 5. The thin curve, on the other hand, represents the case in which $\alpha = \beta = 0$. The critical punishment probability that makes every individual indifferent between work and evading taxes is then given by

$$\tilde{\pi} = \frac{\ln(1-\rho\tau) - \ln(1-\tau)}{\ln(1-\rho\tau) - \ln(1-\tau) - \ln(1-\gamma)} \simeq 0.4. \tag{32}$$

Nobody evades taxes if $\pi > \tilde{\pi}$ while everyone does so if $\pi < \tilde{\pi}$.

5.4. Figure 6. The numerical specification behind this diagram is $\bar{w} = 1$, $C_0 + E = c = 0.1$, and $\pi = 0.5$. The resulting political equilibrium condition (29) then becomes

$$u \geq -1 \ln\left(1 - 0.1\frac{1+x}{1-x}\right)$$

or, equivalently,

$$u \geq \ln\left(\frac{1-x}{0.9-1.1x}\right).$$

References

[1] Allingham, M. and A. Sandmo. "Income tax evasion: a theoretical analysis". *Journal of Public Economics*, Vol.1, 1972.

[2] Becker, G.: "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 1968.

[3] Bentham, J. *An Introduction to the Principles of Morals and Legislation*". Chapter XII, Section I. 1789.

[4] Bar-Gill, O. and A. Harel. "Crime Rates and Expected Sanctions: the economics of deterrence revisited". *Journal of Legal Studies.* 2001.

[5] Elster, J. "Social Norms and Economic Perspectives", *Journal of Economic Perspectives.* 1989.

[6] Fender, J. "A General Equilibrium Model of Crime and Punishment". *Journal of Economic Behavior & Organization.* 1999.

[7] Frey, B. and L. P. Feld "Deterrence and morale in taxation: An empirical analysis ", CES*ifo Working Paper*, No. 760. 2002

[8] Huck, S., D. Kübler and J. Weibull. “Social Norms and Economic Incentives in Firms.” Working Paper. 2004.

[9] Lindbeck, A, S. Nyberg and J. Weibull. “Social Norms and Economic Incentives in the Welfare State.” *Quarterly Journal of Economics.* 1999.

[10] Lindbeck, A, S. Nyberg and J. Weibull. “Social Norms and Welfare State Dynamics.” *Journal of the European Economic Association.* 2003.

[11] Pommerehne, W.W. and H. Weck-Hannemann "Tax rates, tax administration and income tax evasion in Switzerland", *Public Choice*, Vol. 88. 1996

[12] Rasmusen, E. “Stigma and Self-fulfilling Expectations of Criminality”. *Journal of Law and Economics.* 1996.

[13] Sandmo, A. "The Theory of Tax Evasion: A retrospective view ". *Working paper.* 2004