Centre de recherche sur l'emploi et les fluctuation économiques (CREFÉ)
Center for research on Economic Fluctuation and Employment (CREFE)

Université du Québec à Montréal

Cahier de recherche/Working Paper No. 62

Public Safety, Altruism and Redistribution*

Dominique Demougin
Otto-von-Guericke University Magdeburg, Germany and CREFE

Robert Schwager
Otto-von-Guericke University Magdeburg, Germany

May 1998

Demougin: Phone number: (+49) 391 67-18545, E-mail: demougin@ww.uni-magdeburg,de
Schwager: Phone number: (+49) 391 67-18518, E-mail: schwager@ww.uni-magdeburg.de

* The authors are grateful to Claude Fluet and the participants of the 1998 conference on Public Economic theory in tuscaloosa, USA, for insightful comments. We would like to thank two anonymous referees for useful comments. We retain full responsibility for any error.

Résumé. Un modèle est développé dans lequel une partie de la population poursuit des activités illégales. Les incitations à la criminalité dépendent de la probabilité et des coûts d'opportunité d'être puni. Le gouvernement peut influencer ces variables par un système de coercition et par sa politique sociale. Nous étudions d'une part les choix politiques qui minimisent les coûts et d'autre part, les incitation de la population "riche" à financer une politique de redistribution. Nous comparons cette approche avec d'autres modèles où les politiques de redistribution résultent des préférences altruistes.

Abstract. A model where a portion of the population participates in illegal activities is developed. The propensity to participate in crimes depends on the probability and the opportunity cost of being arrested. The government can influence these variables by law enforcement expenditure and by social transfer. We analyze the cost minimizing policy mix and the incentive of wealthy agents to finance redistribution. We compare this approach with standard modelling where redistribution follows from altruistic preferences.

JEL Classification: H8, K0

Keywords: Criminality, Redistribution, Public Safety

In the Western world, some provision for those threatened by the extremes of indigence or starvation due to circumstances beyond their control has long been accepted as a duty of the community. ... The necessity of some such arrangement in an industrial society is unquestioned – be it only in the interest of those who require protection against acts of desperation on the part of the needy.

F. Hayek

1. Introduction. In the present paper, we examine the need for security as *raison d'être* for the introduction of distributional objectives. We find that safety considerations on the part of tax payers do, indeed, warrant altruistic like preferences. However, we also show using an example that the resulting distributional trade-offs cannot always be adequately captured by a welfare function.

To substantiate our questioning consider one of the standard arguments advanced to theoretically justify income maintenance policies. "The rich are assumed to be altruistic toward the poor and, therefore, to be willing to incur tax liabilities to support redistributive transfers to the poor"[1] (Wildasin (1991)). This altruistic motive is then formalized in a welfare function which pins down a specific marginal rate of substitution between the income of wealthy and poor agents. What factors determine this marginal rate of substitution? How is this rate affected by variations in the underlying economic structure? We provide one possible answer to these questions by modelling an altruistic like behavior founded on tax payers' demand for safety. In our model, the rich consider and may care about the income level of the poor because transfers are one instrument to produce the public good of safety.[2]

Explicitly, we model the incentives of potential criminals. We consider a world with two types of individuals distinguished by their productivity. High productivity agents pay taxes and may also become victim of a crime. Low productivity agents could receive welfare payments and may decide to become thieves. The incentive to get involved in such a crime will depend on the probability and the opportunity costs of being arrested. Governments can influence both variables by spending more or less money on law enforcement[3], and by varying the level of welfare benefits. In a first result, we derive the cost minimizing policy mix for an exogenously given level of public safety. We find that for a high level of criminality, it is cost effective to exclusively rely on law enforcement, and to pay no transfer. However, we also show that in order to achieve a low level of criminality, it will at some point become advantageous to redistribute income. Heuristically, transfers can only be effective if there is a possibility to lose them when an agent undertakes a crime. Hence, one must

[1] Wildasin (1991) p. 757.

[2] See the above quotation from Hayek (1960, p. 285). Heuristically, the argument can also be found in Pauly (1973, p. 38) who points out that "if poverty contributes to the incidence of crimes against property and persons, one way to reduce crime may be to redistribute income". For an empirical argument, see Wong (1995).

[3] We abstract from all considerations of private provision of security because of the well-known public good nature of safety.

first rely on law enforcement. At this stage, whether or not altruistic like behavior occurs, depends on the desired level of safety.

However, safety itself is an endogenous variable that should be determined within the model. Following standard practice, we analyze this choice from the point of view of those who finance transfers.[4] They trade off the damage of criminality against the costs of security. These costs are given by the tax required to finance the policy mix along the expansion path and the associated excess burden. In analogy to taxes, we define the excess burden of crime as the difference between the gain to the offender and the damage to the victim. We find that this difference plays a crucial role in the analysis. When there is no such excess burden of criminality, redistribution through the tax system is never optimal.[5] To use a metaphor, we are in an environment where taking the medicine would make the patient worse off than the actual illness. Specifically, to reduce the cost of criminality by 1$ using transfers would always impose a burden of more than 1$, and this for two reasons. First, because taxation causes a deadweight loss. And second, because social security cannot discriminate between potential criminals and other poor individuals. As a result, to reduce criminality via a tax-transfer scheme all the unproductive agents would have to be 'bribed'. Consequently, in the absence of an excess burden of crimes altruistic like behavior cannot be founded on safety considerations.

However, a deadweight loss of crimes appears very reasonable. An obvious justification is that stolen goods cannot be sold at market prices. Further examples are the possibility of physical pain inflicted by an injury, the psychological suffering or even risk aversion concerns on the part of victims . If one of these excess burdens is large enough, it becomes optimal even for the wealthy to advocate a redistribution of income in order to increase public safety. This shows that, indeed, the presence of an excess burden of crime can justify altruistic like behavior. However, the resulting trade-offs between the income of the wealthy and that of the poor is not fixed by a true altruistic preference structure. Instead it reflects cost minimizing behavior of the rich where the occupational choice of the unproductive agents enters as a constraint. Consequently, any changes in the underlying parameters, like e.g. the excess burden of crime, will affect the trade off.

Our paper is closely related to two strands of the literature. Firstly, many other reasons have been explored why transfers to the poor may be in the interest of taxpayers. For example, redistribution may emerge from a social contract concluded under the veil of ignorance (Rawls (1972)), or it may provide an insurance against income shocks occurring during the life cycle (Varian (1980)). In this kind of argument, those who end up rich might want to avoid taxation *ex post* but redistribution increases their utility *ex ante*. Moreover, Atkinson (1995, p. 2) claims that income maintenance policies may induce the recipients to take more risk, and Grüner (1998) shows that

[4]For example, Wildasin (1991, p. 763) assumes that "redistributive transfer policy is set in such a way as to maximize the welfare of the rich". A similar approach is used by Pauly (1973).

[5]Thinking of Robin Hood in Sherwood Forest, it should be noted that criminality is also a form of redistribution. However, unlike the tax system, it imposes a lump-sum burden.

transfer payments may improve the allocation of venture capital since able but poor entrepreneurs become more credit-worthy. Such an improvement in the allocation of risk presumably is also beneficial to those who finance transfers. Finally, genuine altruism cannot be ignored. Indeed, its existence can be explained by evolutionary models (see Stark (1995), chapter 6). Our approach does not contradict any of the arguments based on these factors. It shows, however, that fighting criminality in itself can be a sufficient motivation for redistribution, albeit not the only one. Secondly, our paper is also related to the literature on the economics of crime where, following Becker (1968), the participation in illegal activities is modelled as a rational occupational choice of utility maximizing individuals. As in Furlong (1987), we integrate this activity decision into an equilibrium model.[6] However, Furlong (1987) does not examine the effect of redistribution on the choice of criminal activity. Our paper is also closely related to the recent works by Wong (1995) and Zhang (1997) who address some similar issues, though mainly from an empirical perspective. Using British time series and US cross section data, these authors confirm our main argument: the crime rate decreases if the standard of living of poor persons is increased.

The remaining of the paper is organized as follows. In the next section, we present the model. In section 3, we analyze the trade off between law enforcement and redistribution for a given level of criminality. This generates the cost function of public safety. In section 4, we provide our main result. Section 5 offers some concluding remarks.

2. The Model. We analyze a model with a public sector and two types of agents denoted by $\alpha \in \{0, a\}$, where α indicates the productivity of each of the players. The proportions of types within the population are exogenously fixed and given by θ and $1-\theta$ respectively. In our model, the only role of the public sector is to collect taxes in order to finance law enforcement (hereafter, we refer to this public activity as police) and redistribute income.

2.1. Legal Activity. For simplicity low productivity agents (type $\alpha = 0$) are assumed to have no labor income at all. However, through a welfare system non-working agents may receive a lump sum transfer $\tau \geqslant 0$. For parsimony, we assume that all the agents have the same preferences which can be represented by a utility function that is linear in revenue. The utility of a law abiding type-0-agent becomes:

$$v_0(\tau) = \tau. \tag{1}$$

An agent with high productivity receives a positive income net of her effort costs which is denoted by a. However, in order to finance police and transfers the productive participants must be taxed. In order to raise T dollars of tax revenue from a type-a-agent, it is necessary to impose on her a utility loss of $k(T) \geq T$. The function $k(T)$ incorporates both the actual tax payment T and an excess burden of taxation. Tax

[6] For other related references, see e.g. Ehrlich (1981), Usher (1986).

revenue is assumed to be bounded above by $\bar{T}$, reflecting the maximum of the Laffer curve. This interpretation of k imposes some natural restrictions: $k(0) = 0$, $k'(0) = 1$, $k'(\bar{T}) = \infty$ and $k''(T) > 0$ for all $0 \leqslant T \leqslant \bar{T}$.

2.2. Criminal Activity. Agents may become criminally active. For simplicity, the model is designed in such a way that type-a-agents never choose to do so. This should not be interpreted as a better moral on the part of high productivity players. Rather it is thought to reflect different incentives. Specifically, a is assumed to be large enough so that the opportunity cost of undertaking a criminal activity becomes prohibitive for productive agents.[7] The same argument also guarantees that it does not pay for a type-a-agent to stop working in order to draw from the social security system.

Criminals are assumed to rob only once and either obtain x, $0 < x \leqslant a - k(T)$, if they encounter a type-a-agent or 0 otherwise. We denote by p_0 the conditional probability that a type-0-agent becomes criminal. Taking into consideration that only a low productivity agent will find it optimal to engage in a criminal activity, the probability that a randomly chosen agent is a thief becomes $p = \theta p_0$. Agents are assumed to encounter only one other person at random. Thus, p also denotes the probability that an $\alpha = a$ agent will be robbed.

While x gives the return of a successful crime, we allow for an excess damage imposed on the victim. Specifically, a type-a-agent, who may be robbed with the probability p, sees her expected utility reduced by $d(p) \geqslant px$. The excess burden, measured by the difference $d(p) - px$, can be justified in numerous ways. For example, we would expect that a criminal cannot sell stolen goods at market prices. A further justification is the possibility of non-pecuniary damages to the victim due to pain inflicted by an injury or more generally to risk aversion considerations. The function d satisfies $d'(p) \geqslant x$ and $d''(p) \geqslant 0$ for all $0 \leqslant p \leqslant \theta$. Altogether, the utility of a productive agent becomes

$$v_a(T, p) = a - k(T) - d(p). \tag{2}$$

2.3. Occupational choice. The decision process of a type-0-agent is summarized in figure 1. Since an offender chooses a victim at random, his expected gain from crime becomes $(1 - \theta)x$. Thus, conditional on not being arrested, a thief will enjoy the expected utility $v_{NA} = (1 - \theta)x + \tau$. If arrested, the criminal is assumed to lose his entire monetary assets consisting of x and τ. For simplicity, we assume that these values are dissipated rather than confiscated. In addition, an arrested criminal suffers a nonmonetary utility loss from being sentenced to prison with the

[7] A sufficient condition for this is:

$$a > k(\overline{T}) + \frac{1-\theta}{\theta}\overline{T} + (1-\theta)x + d(\theta).$$

Our main results would not be affected if we allowed a proportion of type-a-agents to become criminally active.

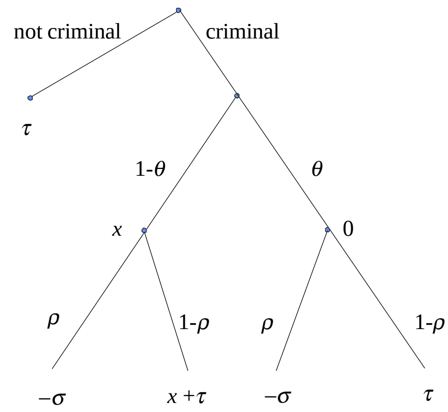

Figure 1: Decision tree of a type-0-agent

monetary equivalent measured by σ. Hence, conditional on being arrested $v_A = -\sigma$ is the expected utility of an offender. The disutility of imprisonment is assumed to be a random variable reflecting different preferences on the part of the agents. For parsimony, σ is assumed to be uniformly distributed[8] on the interval $[0, s]$.

From the point of view of a criminal, ρ denotes his probability of being arrested and found guilty through law enforcement efforts. The cost of inducing ρ is $c(\rho)$ per capita of the unproductive part of the population, hence $\theta c(\rho)$ per capita of the overall population.[9] We assume $c(0) = c'(0) = 0$, for $\rho > 0$, c', $c'' > 0$, and $c'(1) = +\infty$. Thus, a criminal will enjoy v_{NA} with probability $(1 - \rho)$ and v_A with probability ρ. Hence, his overall expected utility becomes

$$v_c = (1 - \rho)\left[(1 - \theta)x + \tau\right] - \rho\sigma. \tag{3}$$

A type $\alpha = 0$ agent will choose to become criminally active if $v_c > v_0$, i.e.

$$(1 - \rho)(1 - \theta)x - \rho\tau > \rho\sigma. \tag{4}$$

Because of the assumption of a uniform distribution for σ, we conclude that

$$p_0 = \begin{cases} 1 & \text{if} \quad \frac{\left(\frac{1}{\rho}-1\right)(1-\theta)x-\tau}{s} > 1 \\ \frac{\left(\frac{1}{\rho}-1\right)(1-\theta)x-\tau}{s} & \quad \text{otherwise} \\ 0 & \text{if} \quad \frac{\left(\frac{1}{\rho}-1\right)(1-\theta)x-\tau}{s} < 0. \end{cases} \tag{5}$$

[8]From the point of view of our main conclusions, the assumption of a uniform distribution is inessential. We use it because of its simplicity and because it avoids second-order derivatives. Its main weakness is that the support must be bounded.

[9]In this modelization, the per capita cost of ρ is assumed to be independent from the actual size of the criminal sector. Heuristically, police is presumed to control persons and not investigate crimes. An example of this may be the monitoring of income tax returns where a given proportion of returns are verified.

From this equation, it is immediately obvious that p_0 is decreasing in ρ and τ. This reflects that in our model police activity and welfare payments are two policy instruments that may be used to reduce crime. Law enforcement increases the risk and, thus, the expected cost of unlawful behavior, whereas redistribution[10] increases the expected opportunity cost of losing transfers. As can be seen from the equation, this second effect can only occur when there is some level of law enforcement since otherwise transfers can never be lost.

3. The cost of public safety.

In this section, we keep the level of public safety – measured by p_0 – constant and minimize expenditures. We derive the cost function of crime deterrence and the optimal instrument mix. The resulting conditions are necessary for any policy to maximize welfare. Moreover, the analysis is interesting in its own right because of the dual. Indeed, in many countries lower level governments have no discretion over their tax revenues, but have a choice to allocate funds to different expenditure categories.

3.1. The Government Budget.

The government budget constraint is given in per capita terms. Since only type-a-agents pay taxes revenue is $(1-\theta)T$. From the foregoing section, the police expenditures are $\theta c(\rho)$. Finally, transfers are only paid to low productivity agents, thus, yielding a total volume of transfer expenditures of $\theta\tau$. Altogether, the government budget constraint becomes

$$(1-\theta)T - \theta c(\rho) - \theta\tau = 0. \tag{6}$$

For every tax T, the government budget equation yields an isoexpenditure curve. As can be seen from figure 2, it generates in the (ρ, τ) space a strictly concave budget constraint defined by

$$\tau = \frac{1-\theta}{\theta}T - c(\rho) \tag{7}$$

with the slope

$$\frac{d\tau}{d\rho} = -c'(\rho) \leqslant 0 \tag{8}$$

which is zero at $\rho = 0$ and strictly negative otherwise. Higher revenues T generate a parallel upward shift of the budget constraint.

3.2. Isoprobability Curves.

The definition of p_0 in (5) yields a map of strictly convex isoprobability-curves represented in figure 3:

$$\tau = \frac{(1-\theta)x}{\rho} - [sp_0 + (1-\theta)x] \tag{9}$$

with slope

$$\frac{d\tau}{d\rho} = -\frac{(1-\theta)x}{\rho^2} < 0. \tag{10}$$

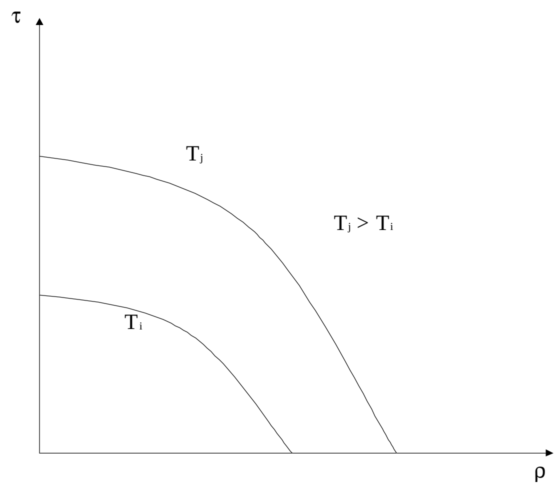

Figure 2: Isoexpenditure Curves

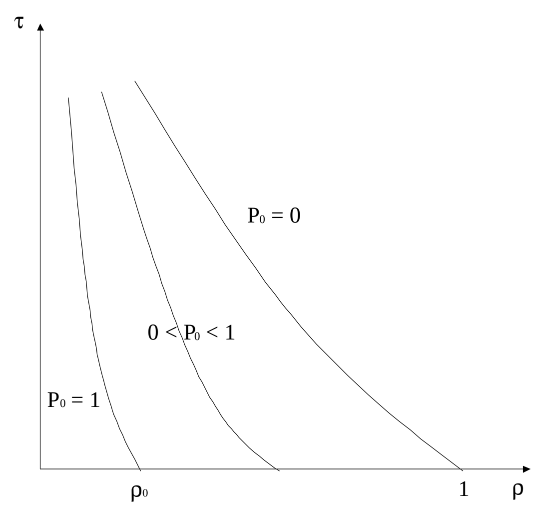

Figure 3: Isoprobability Curves

Geometrically, reducing p_0 shifts the isoprobability curve outward, the intuition being that in order to reduce crime the government must either increase transfers or law enforcement or both. For $\rho = 1$, the probability of crime drops to zero even without any transfer. This is a natural property in a model with rational agents: nobody will choose to become criminally active if the probability of being arrested is 1. Hence the $p_0 = 0$-curve cuts the horizontal axis at $\rho = 1$. At the other extreme, the $p_0 = 1$-curve cuts the axis at

$$\rho_0 = \frac{(1-\theta)\,x}{s + (1-\theta)\,x}. \tag{11}$$

Heuristically, ρ_0 denotes the highest arrest probability such that when there are no transfers all the type-0-agents choose to become thieves . From the equation, it appears reasonable to think of s as being large relative to $(1-\theta)x$ (i.e. ρ_0 is small). Specifically, we assume $(1-\theta)x < s$.[11] In figure 3 the area to the left of the curve going through point ρ_0 represents the set of policies which all lead to $p_0 = 1$. Further, policies depicted in the area above the curve going through the point $\rho = 1$ induce zero criminality.

3.3. Optimal expenditure mix. In this subsection, we define the cost function of public safety, $C(p_0)$, as the minimal expenditure necessary to keep criminality[12] to an exogenously predetermined level $p = \theta p_0$. From the minimization problem, we further obtain the optimal instrument mix. We use the definition of p_0 from equation (5). It is easily seen that for $p_0 = 1$ the cost minimizing policy mix yields $T = \rho = \tau = 0$. The intuition should be obvious: if one is ready to accept the maximal possible level of criminality one will not engage in any public safety expenditures. For all other levels of public safety, $p_0 \in [0, 1)$, we solve:

$$C(p_0) = \min_{\tau,\rho} \quad \frac{1-\theta}{\theta}T = c(\rho) + \tau \tag{I}$$

$$p_0 \geqslant \frac{(\rho^{-1}-1)(1-\theta)x - \tau}{s} \tag{12}$$

$$\tau \geqslant 0. \tag{13}$$

The first inequality states that criminality must not exceed the predetermined level p_0. Geometrically, it implies that the set of permissible policies are to the north-east

[10] This redistribution effect formalizes the argument of Hayek (1960) and Pauly (1973) cited earlier.

[11] Thus, $\rho_0 < 1/2$. In reality s is also controlled by society. Using the well known argument by Becker (1968), we can assume that punishment is maximal, so that the assumption is not very restrictive. The assumption is used only once, in the analysis of the case without any excess burden of criminality, see proposition 3.

[12] To define $C(p_0)$ as public safety is a marginal abuse of language, since it actually denotes the public costs necessary to keep criminality at $p = \theta p_0$.

corner of the p_0-isoprobability curve. The second constraint guarantees that transfers cannot become negative, i.e. that poor agents cannot be taxed. In order to solve the above minimization problem, we write the Lagrangian

$$\mathcal{L} = c(\rho) + \tau + \lambda \left(p_0 - \frac{(\rho^{-1} - 1)(1-\theta)x - \tau}{s} \right) + \mu\tau \tag{14}$$

which yields two first-order conditions

$$\mathcal{L}_\rho = c'(\rho) + \lambda \frac{(1-\theta)x}{s\rho^2} = 0 \tag{15}$$

$$\mathcal{L}_\tau = 1 + \lambda \frac{1}{s} + \mu = 0 \tag{16}$$

and two complementary slackness conditions

$$\lambda \left(p_0 - \frac{(\rho^{-1} - 1)(1-\theta)x - \tau}{s} \right) = 0 \tag{17}$$

$$\mu\tau = 0 \tag{18}$$

with $\lambda, \mu \leqslant 0$ since we have a minimization problem. Using these equations, one can characterize the solution to problem (I).

First, we observe that $\lambda < 0$. Indeed, if $\lambda = 0$ were feasible, then from (15) $c'(\rho) = 0$. By assumption this would imply $\rho = 0$, hence $\frac{(\rho^{-1}-1)(1-\theta)x-\tau}{s} = \infty$ which violates (12). Intuitively $\lambda = 0$ is only feasible at the point $\rho = \tau = 0$ which has been ruled out.

From the second multiplier, we have two possible cases.

1. $\mu = 0$. In that case, we have an interior solution with

$$c'(\rho) = \frac{(1-\theta)x}{\rho^2} \tag{19}$$

   and τ follows implicitly by (17).

2. $\mu < 0$. In that case, we have a corner solution with $\tau = 0$ and ρ follows implicitly by (17).

The optimal policy mix is summarized in the following proposition.

PROPOSITION 1. *Define $\widehat{\rho}$ as the solution of (19). For $p_0 = 1$, we have $\rho^* = \tau^* = 0$. For $p_0 \in [0, 1)$ we obtain*

1. *if $p_0 \leqslant (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$ then $\rho^* = \widehat{\rho}$ and τ^* is implicitly defined by (17).*
2. *if $p_0 > (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$ then $\tau^* = 0$ and ρ^* is implicitly defined by (17).*

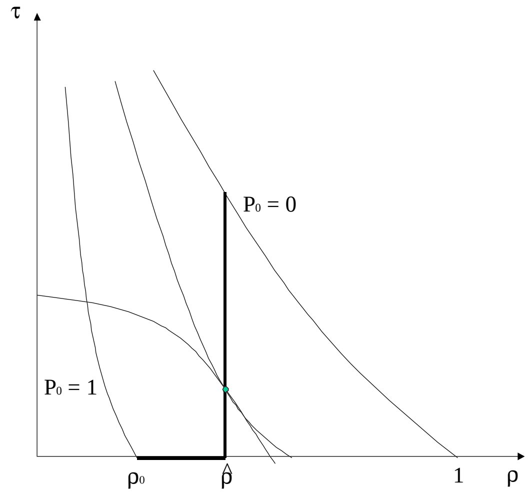

Figure 4: Expansion path

At an optimal policy mix where both instruments are used (case 1), ρ^* is defined by equation (19) which equalizes the marginal rate of transformation between social security and law enforcement along a budget constraint (the left handside of the equation) to the marginal rate of substitution along the isoprobability curve. Since both of these rates only depend on the arrest probability ρ, it must remain constant at $\rho^* = \widehat{\rho}$ as long as one has an interior solution.

In the case of a corner solution (case 2), it can be seen from the first-order conditions of the Lagrangian that in absolute value terms the marginal rate of substitution along the isoprobability curve is larger than the marginal rate of transformation. Were we not at a corner solution, the government would want to substitute transfers for law enforcement. An alternative way to restate this observation is that if government could tax unproductive agents ($\tau < 0$), it would marginally do so and spend the revenue generated on police.

Figure 4 gives the expansion path which results when we vary the level of criminality p_0. For the no criminality case, i.e. $p_0 = 0$, we argue that a corner solution ($\mu < 0$) is not possible. This follows immediately from the foregoing paragraph because at the point $\rho = 1, \tau = 0$ the marginal rate of substitution along the isoprobability curve is bounded but the marginal rate of transformation $c'(1)$ is infinite. By continuity, the argument extends to p_0 close to zero. For small shifts, we remain in an interior solution. Thus, police expenditure remains constant with $\rho^* = \widehat{\rho}$ and an increase in p_0 implies a reduction in transfers. This downward movement along the expansion path is bounded by one of two situations. Either p_0 becomes 1 or we reach a corner solution[13] — as in figure 4 — where transfers are zero. Obviously, once this point is

[13] In figure 3, we have represented the case where for some p_0 a corner solution results. This will occur if and only if $\widehat{\rho} \geq \rho_0$. This will only happen if at the point ρ_0 a corner solution is optimal, i.e.

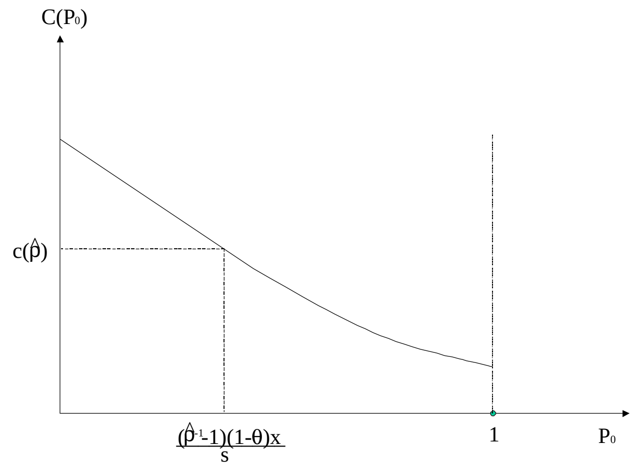

Figure 5: Cost function of public safety

reached any further increase in criminality can only be accommodated by a reduction in law enforcement.

As a result from the optimization, we obtain the function $C(p_0)$ which denotes the cost of keeping criminality at the level p_0. From the envelope theorem, we know that $C'(p_0) = \lambda = -s(1+\mu)$, where the second equality follows from (16). We have shown earlier that $\lambda < 0$, thus, costs are strictly decreasing in criminality or equivalently increasing in public safety. As long as we are in an interior solution, the marginal cost in p_0 is $-s$. Geometrically, we are on the vertical part of the expansion path and any variation in public safety is entirely absorbed by an appropriate variation in transfers. As a result, marginal cost follows from the partial $\partial p_0/\partial\tau$. In the case of the corner solution, the government cannot keep the marginal costs of both instruments equalized. Therefore a further increase in p_0 translates in less cost saving, i.e. $-C'(p_0) < s$. Results are summarized in figure 5.

The last two figures are also of interest for the dual problem where a local government maximizes public safety for an exogenously given level of income. Given a level of revenue, we can read from figure 5 the optimal p_0 and from figure 4 the resulting policy mix.

4. When would the rich desire redistribution?. In this section we complete the analysis by solving the trade-off between taxes and public safety. We do this from the perspective of a representative citizen of type a. This approach is motivated by our main question: does the demand for public safety induce the taxpayers to

---

if

$$c'(\rho_0) \leqslant \frac{(1-\theta)x}{\rho_0^2}.$$

behave as if they were altruistic? Specifically, we ask under what circumstances does it become advantageous for those who finance the transfers to redistribute income.[14]

From the foregoing section, we use the cost function for public safety to define the indirect utility of a representative a-type agent:

$$\max_{T,p_0} v_a(T,\theta p_0) = a - k(T) - d(\theta p_0) \tag{II}$$

$$C(p_0) = \frac{(1-\theta)}{\theta}T \tag{20}$$

$$0 \leqslant T \leqslant \overline{T} \tag{21}$$

Since $k'(\overline{T}) = \infty$, a solution will never imply $T = \overline{T}$. On the other hand, a corner solution with $T = 0$ is possible because of the discontinuity of $C(\cdot)$ at the point $p_0 = 1$ which corresponds to the maximum level of criminality. From figure 3, we note that the $p_0 = 1$-curve is bounded away from the origin. In other words, starting at $p_0 = 1$, a marginal decrease in criminality requires a discrete jump in taxation. Yet the resulting benefit can only be marginal since all the other curves are differentiable. Thus, there always is a local maximum at $T = 0$.

Though it is analytically feasible to construct cases, depending on the curvature assumptions for $k(\cdot), d(\cdot)$ and $C(\cdot)$, where the no tax solution becomes the global maximum, we do not regard this case as particularly interesting from an empirical as well as from a theoretical perspective. The $T = 0$ case occurs if all the low productivity agents choose to be criminally active so that marginal changes in either form of public expenditure become ineffective. Theoretically, this counterintuitive result is an artifacts of our simplifying assumption that σ is uniformly distributed and therefore has a finite support.[15] From an empirical perspective this case does not appear convincing either. For example, Blank (1996) states that "criminal activity ... are, even in the poorest neighborhood, the activity of a few".[16] For these reasons, in the remainder, we restrict attention to the case where the optimal level of criminality p_0^* is strictly below unity. Consequently, all the multiplicators from (21) are zero.

Substituting the cost of public safety in the objective function of the representative type-a-agent and maximizing with respect to p_0 yields the following condition:[17]

$$-k'\left(\frac{\theta}{(1-\theta)}C(p_0)\right)\frac{\theta}{(1-\theta)}C'(p_0) = \theta d'(\theta p_0). \tag{22}$$

This equation has a straightforward interpretation. It equates the marginal damage of criminality (the right handside) to its marginal benefit from the reduction in required

[14] At this point, it is worthwhile to go back to the issue of private provision of safety which by assumption was ruled out (see footnote 3). Since in our model, all the wealthy agents are identical, their demands for safety will be the same. Thus the choice of public safety by the representative agent would crowd out any private provision of security.

[15] For a discussion of this point, see footnote 8.

[16] Blank (1996) p. 51.

[17] It is easy to derive that the objective function is concave for all $p_0 < 1$. This follows from the slope and the curvature assumptions for $k(\cdot)$ and $d(\cdot)$ and from the observation that $C'' \geqslant 0$.

taxes. This benefit depends itself on the marginal cost of taxes and on the effectiveness of government revenue to affect criminality.

Since minimizing public expenditure is a necessary condition to resolving (II), we can immediately apply proposition 1 to p_0^*. Thus: .

COROLLARY 2. *For* $p_0^* \in [0,1)$, *we obtain:*

1. *if* $p_0^* < (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$ *then* $\tau^* > 0$,
2. *if* $p_0^* \geqslant (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$ *then* $\tau^* = 0$.

In this form the result is not very interesting. All it states is that transfers will be positive if the optimal level of criminality is smaller than the level represented by the isoprobability curve going through the point $\widehat{\rho}$. However, the corollary allows us to compare our approach with the standard modelling where redistribution is derived from optimization of a welfare function. The tradeoff between the incomes of wealthy and poor individuals is not fixed by the form of such a function, but depends on the equilibrium behavior of individuals. Specifically, consider the definition of p_0 in equation (5) which describes the occupational choice of low-productivity agents. In a situation where $p_0^* < (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$, one can insert p_0 into the utility function v_a in (II) and use $\rho^* = \widehat{\rho}$ to compute the marginal rate of substitution between the tax and the transfer:

$$\frac{dT}{d\tau}_{|v_a} = \frac{\theta d'}{sk'} > 0. \tag{23}$$

However, if $p_0^* \geqslant (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$, the wealthy agents will always choose $\tau^* = 0$ and the marginal rate of substitution is

$$\frac{dT}{d\tau}_{|v_a} = 0. \tag{24}$$

Thus the marginal willingness to pay for redistribution does not only depend on the preferences of the wealthy agents. In addition, the marginal damage of criminality (d'), the excess burden of taxes (k'), the incentives faced by potential criminals ($(1-\theta)x$ and s), and the marginal costs of law enforcement ($c'(\rho)$) are important. In particular, these structural variables determine which of the cases of the corollary is relevant, and hence, whether the wealthy agents do indeed behave as if they had altruistic preferences.

In the rest of the section, we provide some intuition for the kind of economic environments which generate such altruistic like behavior. We do this by presenting several specifications for the damage function $d(p)$ which highlight the conditions leading to a positive transfer. First, we consider the case where there is no excess burden of criminality, i.e.

$$d(p) = px. \tag{25}$$

Here, the loss of the victim is exactly as large as the gain for the offender. In this formulation, crime can be thought a 'pure and socially costless redistribution of income'. As we now show, if this were true, it would never pay to use transfers to reduce criminality.

PROPOSITION 3. *If there are no excess burdens of criminality then transfers are always zero.*

Proof: Use $d' = x$, $k' > 1$ and $s > (1-\theta)x$. It follows $p_0^* \geqslant (\widehat{\rho}^{-1} - 1)(1-\theta)x/s$. Q.E.D.

To understand the intuition of this result, let us initially imagine that all the unproductive population became involved with crime. It is easy to see that transfers would never be used. Indeed, as mentioned above, both transfers and stealing can be thought as one way to redistribute income, but transfers are more expensive due to the excess burden of taxes. The argument is even further reinforced when only a fragment of the unproductive population engages in crimes because transfers cannot discriminate and, thus, be targeted to potential offenders. From this we learn that redistribution will only be considered by the rich as a possible instrument when there is an excess burden of criminality. In other words, the existence of a deadweight loss of crimes is a necessary condition for altruistic like behavior to occur.

In the next two results, we examine two possible cases of such an excess burden. In the first extension, we consider situations where offenders cannot resell stolen goods at market prices. As a result, the damage to a victim is larger than the gain to the thief. We denote with y, $y > x$, the loss to the victim which yields

$$d(p) = py. \tag{26}$$

With this formulation, a transfer is optimal if y is large enough. Specifically

PROPOSITION 4. *Define* $\widehat{y} = \frac{s}{1-\theta}k'\left(\frac{\theta}{1-\theta}c(\widehat{\rho})\right)$, *then if* $d(p) = py$, *for all* $y > \widehat{y}$, *the optimal transfer is strictly positive.*

The proof follows immediately from (22) and is left for the reader. Heuristically, variations in y leave the marginal benefit of criminality (the left handside of (22)) unchanged. However, the marginal damages of criminality increase linearly in y reducing the optimal level of criminality p_0^*. At some point, p_0^* will attain the critical value $(\widehat{\rho}^{-1} - 1)(1-\theta)x/s$. Any further increase in y will move the optimal policy along the vertical part of the expansion path and transfers become positive.

Finally, we consider a case where the excess burden is due to risk aversion considerations on the part of the type-a-agents. For simplicity, we assume a mean-variance-preference structure. We abstract from other kinds of excess burden so that the expected loss is px and the variance $p(1-p)x^2$. Thus, we may write

$$d(p) = px + r \cdot p(1-p)x^2, \tag{27}$$

where $r > 0$ measures the degree of risk aversion. As long as the probability of being a victim of a crime is less than one half, reducing p will not only reduce the expected loss but also the variance. This increases the marginal cost of a crime and hence makes the optimal tax larger, and the optimal p_0 smaller. When risk aversion becomes large enough, again the optimal level of criminality falls below the critical value $(\widehat{\rho}^{-1} - 1)(1-\theta)x/s$, and the optimal transfer must be strictly positive.

Proposition 5. *Define $\widehat{r}$ as the implicit solution of*

$$\frac{s}{1-\theta}k'\left(\frac{\theta}{1-\theta}c(\widehat{\rho})\right) - \left[x + r\left(1 - 2\theta\frac{(\widehat{\rho}^{-1}-1)(1-\theta)x}{s}\right)x^2\right] = 0 \qquad (28)$$

then if $d(p) = px + r \cdot p(1-p)x^2$ for all $r > \widehat{r}$, the optimal transfer is strictly positive.

To summarize, transfers are the more likely to be used as an instrument to reduce criminality, the higher the wedge is between the benefit of crime for the criminal and the cost of crime for the victim. If this deadweight loss of criminality is large enough, altruistic like behavior can indeed be founded on the demand for public safety on the part of taxpayers.

5. Conclusion. In this paper, we have shown that redistribution can in some cases be in the interest of those who finance transfers because it promotes public safety. From a cost minimizing perspective, we found that transfers across income groups are only optimal if a high level of public safety is to be achieved. Otherwise our model suggests that pure law enforcement minimizes expenditures. We also derived the optimal amount of taxes which wealthy agents are willing to pay in order to reduce criminality. We found that, when there are no excess burdens of criminality, the level of taxes chosen is so low that no transfers are ever paid. However, the assumption of no excess burden does not appear very realistic. We constructed two simple cases where an excess burden will lead the wealthy agents, for some parameters, to find some level of redistribution optimal. Intuitively, this occurs when the excess burden is large. Thus, our paper showed that safety concerns lead to the same behavior as altruistic preferences if and only if criminality creates a substantial deadweight loss.

Our paper yields similar results to models which derive redistribution from a welfare function representing altruistic preferences of taxpayers. It therefore could be thought as a possible justification of that approach. However, our analysis also shows that, and how, the incentives for wealthy agents to redistribute income depend on the optimizing behavior of potential criminals. When using altruistic like preferences in applied models, one therefore cannot fix a functional form for a welfare function without taking this behavior into account.

A prominent example of such an application is migration. Based on our model, one might analyze how the willingness to pay for redistribution changes if transfer recipients are mobile and policies are determined locally. Will the optimal mix of law enforcement and transfers change? Is there a tendency toward lower redistribution as in most models which use altruistic preferences? On the other hand, if taxpayers are mobile, is it true that they necessarily move to low-redistribution jurisdictions, or will the Tiebout mechanism work even in the presence of redistribution? This seems to be likely since in a low tax country, taxpayers also face a higher level of criminality. Redistribution is just an instrument to provide the public good of safety, and therefore should be supplied just as efficiently by local jurisdictions as any other local public good.

Another interesting application of our model would be to explain the widespread use of in kind transfers. We know from the standard model in microeconomic theory that cash transfers should always be better unless one is ready to accept altruistic and paternalistic preferences for the rich. However, since in our approach it is not the utility level of the poor but their behavior that motivates the rich to provide transfers, transfers in kind will be preferable if they have a stronger incentive effect to reduce criminality.

References

Atkinson, A. (1995): *Incomes and the welfare state: Essays on Britain and Europe*, Cambridge University Press.

Becker, G. (1968): "Crime and punishment: An economic approach", *Journal of Political Economy* 76:169-217.

Blank, R. (1996): *It takes a nation: A new agenda for fighting poverty,* Princeton University Press, Princeton.

Ehrlich, I. (1981): "On the usefulness of controlling individuals: An economic analysis of rehabilitation, incapacitation and deterrence", *American Economic Review* 71:307-322.

Furlong, W. (1987): "A general equilibrium model of crime commission and prevention", *Journal of Public Economics* 34:78-103.

Grüner, H. (1998): "Redistribution as a selection device", *mimeo*, University of Bonn.

Hayek, F. (1960): *The constitution of liberty,* Routledge, London.

Pauly, M. (1973): "Income redistribution as a local public good", *Journal of Public Economics* 2:35-58.

Rawls, J. (1972): *A theory of justice*, Oxford University Press, Oxford

Stark, O. (1995): *Altruism and beyond*, Cambridge University Press, Cambridge

Usher, D. (1986): "Police, punishment and public goods", *Public Finance* 41:96-115.

Varian, H. (1980): "Redistributive taxation as social insurance", *Journal of Public Economics* 14:49-68

Wong, Y. (1995): "An economic analysis of the crime rate in England and Wales, 1857-92", *Economica* 62:235-246.

Wildasin, D. (1991): "Income redistribution in a common labor market", *American Economic Review* 81:757-774.

Zhang, J. (1997): "The effect of welfare programs on criminal behavior: A theoretical and empirical analysis", *Economic Inquiry* 35:120-137.