

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
27 Hillhouse Avenue
New Haven, Connecticut 06520

CENTER DISCUSSION PAPER NO. 666

PUBLIC WELFARE AND GROWTH

Xavier Sala-i-Martin

Yale University

June 1992

Note: Center Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments. This is a revised version of the paper originally written March 1992.

Research for this paper was partially funded by the Instituto de Estudios Fiscales in Madrid.

I have benefited from conversations with Robert Barro, Paul Cashin, Larry Katz, Elvinticinc Dedesembre, Fumfum Fum, Chris Sims, Etsuro Shioji, T.N. Srinivasan and Joel Waldfogel.

Abstract

In this paper I develop a simple model of optimal criminal behavior to analyze the role of public welfare policies such as redistributional transfers or wage subsidies. I show that public welfare acts as a crime-preventing device since it increases the opportunity cost of committing crimes.

I argue that transfers and wage subsidies can be thought of as productive public goods subject to congestion, as with police protection and national defense. Transfers and Wage subsidies are productive because they reduce the criminal-induced aggregate distortions in the economy. They are subject to congestion because when a person decides to increase his output he also increases the average output in the economy and, therefore, the reward to others of criminal actions. Hence, he congests the protective role of the public welfare system, for a given level of public welfare payments.

I find the growth-maximizing size of the public welfare program and I show that public welfare should be financed with income (not lump-sum) taxes, despite the fact that income taxes are distortionary. The reason is that income taxes act as a user fee on congested public goods.

Finally, I show that, in a cross-section of 75 countries, the partial correlation between transfers and growth is significantly positive.

A substantial fraction of the recent growth literature deals with the role of government in the process of economic development. Chamley (1981), Lucas (1990) and the subsequent literature deal with the problem of optimal taxation. Barro (1990) developed a model where the productive aspects of public spending are offset by distortionary taxes. He also discusses the optimal size of the government. A number of papers have followed and extended Barro to include public investment, public consumption and different types of other public goods and taxes.

Despite their large and growing size in almost all countries in the world, little attention has been paid to the role of *transfers* and other forms of *public welfare*. In the United States for instance transfers in 1991 represented 12% of GDP and accounted for 46% of spending by the federal government. By comparison, public investment (on which most of the theoretical attention has focused) represents 4% of GDP or 13% of federal spending and national defense is 5.6% of GDP and 21% of federal spending.

In this paper I provide a rationale for the existence of public welfare programs such as redistributional transfers or wage subsidies,¹ and their relation to economic growth. In the existing growth literature, transfers are often modeled as something that enters the utility function of some planner, politician, or median voter (see for instance Alesina and Rodrik (1992), Persson and Tabellini (1991), or Tabellini (1992)), or as a resource which needs to be redistributed in a lump-sum fashion among the population in order to be able to analyze the effect of a particular tax, while keeping constant the overall size of the government (see for instance Blanchard and Fischer (1991)).

The main point of this paper is that transfers are a means to buy social peace.

¹I analyze old-age public pensions in Sala-i-Martin (1992).

They are a way to bribe poor people out of activities that are socially harmful, such as crimes, revolutions, inner-city riots² and other forms of social disruption. In the first section I present a model where people choose the amount of time they want to devote to criminal activities. The model is in the spirit of Becker (1968) and Ehrlich (1970) but, unlike them, I do not try to determine what are the optimal policies to combat illegal behavior (such as the optimal severity of punishment or the optimal size of penalties for different types of crimes). My goal is to show that transfers and other forms of public welfare are devices that reduce the incentive to commit crimes, because they increase the amount of income one can legally receive outside jail. I show that these results are robust to the inclusion of leisure, even though economic intuition says that transfers unrelated to work effort could have a perverse effect on criminal intensity.

One of the key results from the first section is that what matters for criminals is the size of transfers or wage subsidies relative to the average level of income of the economy. Hence, in the aggregate economy, transfers look very much like a public good subject to congestion: when a person increases his income, so does the average income of the economy. This increases the reward to criminal behavior and, with it, the protective role of transfers is congested.

The final section incorporates the analysis into an aggregate model of growth. This allows me to determine the growth-maximizing amount of public welfare in the economy as the government balances the beneficial, protective effects of transfers and wage subsidies with the adverse effects of the distortionary taxes needed to finance such programs. I also show that the government can replicate the planner's solution by using income taxes and not lump sum taxes. The intuition is that income taxes act like 'user fees' on the protective role of the public welfare policies.

²The 1992 Los Angeles riots occurred at exactly the time I was writing the first draft of this paper.

A Simple Model of Criminal Behavior.

The model I use to analyze transfers extends Becker (1968) and Ehrlich (1970). Let t_i be the fraction of time an individual devotes to illegal, criminal or disruptive activities such as thefts, robberies, strikes or revolutions.³ After normalizing total non-leisure disposable time to one, the time devoted to legal activities is $1-t_i$. The reward for devoting one unit of time to a legal activity (work) is the wage rate w . The reward for engaging in criminal activity is βy per unit of time where y is the average income of the economy and β is a number between zero and one. If we argue that criminal activity is akin to mugging people on the street, and that the average person carries a fraction β of his income in his pocket⁴, then $\beta \cdot y$ is the reward per unit of time devoted to crime, and $\beta \cdot y \cdot t_i$ is the reward for criminals who choose to devote t_i units of their time to this activity. I will assume that the only purpose of crime is to obtain the monetary reward. Unlike Becker (1968), agents in my model do not engage in criminal activities simply because they like crime.⁵ Utility is here solely a function of consumption.

Society, through its government, has access to some technology to capture and

³An important and interesting question is why are these activities considered *crimes* that need to be punished, as opposed to just other activities that need to be priced by the market. In other words, a car theft as a transaction between owner and thief. One could think that, instead of being punished, a thief should just pay a price to the owner, just like any other type of transactions. Under this view, there would be no distinction between criminal and non-criminal activities.

⁴This fraction β could be thought of as being chosen by the average person according to some money demand model that I do not need to specify here. It should be noted that, when making this choice, this person will take into account the probability of being mugged and will add it to the interest foregone by holding cash. That is, the larger the number of criminals operating in a certain area, the lower is likely to be the reward per unit of time devoted to crime since people living or working in that area will be careful not to carry too much money in their pockets.

⁵Becker uses this assumption to explain passion crimes and other crimes that entail no direct monetary reward to criminals. Another unrealistic assumption is that all persons in the economy have the same attitude or preference for crime. Different people may perceive crime differently and these differences may be due to educational background and/or religious beliefs.

prosecute criminals. I will assume that the probability of a criminal being caught and convicted is π . This probability should be an increasing function of the effort the government puts into enforcing laws. It could also be thought to be an increasing function of the amount of crime committed by any given person. In this first simple model, however, I will assume that π is independent of the amount of crimes people choose to commit.⁶ I will relax this assumption later on. Stigler (1970) shows that if law enforcement is costly, there is an optimal amount of enforcement which may be lower than the maximum allowed by the current technology. Hence, the probability of capture need not be one, even though achieving such probability may be technologically feasible. We can simply think of π as the probability of capture and conviction given by the existing technology and the optimal level of public effort.

Individual's preferences can be represented by the following expected utility function:

$$(1) \quad U = \pi \cdot \ln(c^P) + (1-\pi) \cdot \ln(c^{nP}),$$

where c^P is the level of consumption if he is caught and convicted (p stands for 'penalized') and c^{nP} is the level of consumption if he is not penalized.

The level of income if he is not convicted is equal to legal work

⁶One could argue that there is learning by doing (or learning by offending): people who commit few crimes are naive and are more likely to be caught. Professional criminals, on the other hand, have more experience and know how evade police more easily. Furthermore, full-time criminals may be able to bribe policemen and judges in order to lower their probability of conviction. The offsetting force is that the more crimes you commit, the more likely the police are to devote their efforts to capture you in particular (while if you are a naive part-time criminal, the police are likely to either ignore you or to spend little effort in trying to capture you). In this simple model I will assume that these forces roughly offset one another and that the probability of being convicted is independent of t_i .

income, $w \cdot (1-t_i)$, plus the income he gets from his criminal activities, $\beta \cdot y \cdot t_i$. I will further assume that there is a **public welfare system** in the economy. Public welfare could take the forms of either a lump-sum transfer T , or a subsidy on the wage, w .⁷ Given that the model is static in nature, all income is consumed so the level of consumption if not convicted is

$$(2) \quad c^{np} = w \cdot (1-t_i) + \beta \cdot y \cdot t_i + T$$

If convicted, individuals must pay a monetary fee, F .⁸ This fee is related to the level of income. This relation could reflect the wages foregone while serving time in jail, or the reduction in lifetime income due to the stigma attached to convicted criminals: conviction may stigmatize offenders by demonstrating that they are untrustworthy. To the extent that jobs that require trust have better wages, the loss of such jobs will be an additional reason why the fee is related to the level of income. Waldfogel (1992) quantifies the importance of this effect empirically. I will

⁷This wage subsidy could take the form of minimum wage laws or the prohibition of work by children (which entails the elimination of the lowest wage jobs).

⁸Some crimes are penalized with physical or non-monetary fees: the death penalty or cutting off the criminal's hands or ears are just two examples (since human ears are not traded in normal markets, these fees should be considered non-monetary). I will, however, abstract from these physical penalties in the present analysis.

Unlike governments, religions try to deter crime by imposing huge non-monetary, yet non-physical, penalties: some would send criminals to hell for infinitely many periods while others threaten criminals with reincarnation into mushrooms or pigs which, unlike cows, are really terrible things to be reincarnated into.

Because these non-monetary fees are paid in future lives, religions need to adopt escape clauses (often called repentment) since, in the absence of such clauses, a person who has committed a crime will find it optimal to keep committing more crimes (once you know you will go to hell or that you will be a pig in a future life, the marginal penalty for committing additional crimes is zero). The problem with allowing for easy repentment is that people may find it optimal to live a criminal life and seek a pardon through last-minute repentment. An interesting calculation would be to determine what is the optimal amount of repentment that should be allowed by religions. I will not try to provide an answer it in this paper.

assume that the fee is homogeneous of degree one in the amount of income one gets if not convicted:

$$(3) \quad F(.) = c^{np} \cdot \lambda \cdot f(t_i),$$

where the fraction of income lost if convicted is $\lambda \cdot f(t_i)$, and $f'(t_i) > 0$, $f(0)=0$ and $f''(t_i) < 0$. In the above, λ indicates the severity of the fee per unit of crime and $f(t_i)$ relates the amount of crime to the severity of the penalty.⁹ The assumption on the concavity of the penalty is made to ensure that the second order conditions are satisfied.¹⁰ Consumption if convicted is therefore

$$(4) \quad \begin{aligned} c^p &= w \cdot (1-t_i) + \beta \cdot y \cdot t_i + T - F(.) = \\ &= (1-\lambda f(t_i)) \cdot [w \cdot (1-t_i) + \beta \cdot y \cdot t_i + T] \end{aligned}$$

One feature of this analysis is that, since the probability of being caught is independent of whether the person actually commits crimes or not, he will have to pay the fee with probability π , even if he sets $t_i=0$ (in other words, people could be

⁹The question of the optimal size of the penalty is interesting and important. For instance, should the penalty be larger than a market compensation to the victim? That is, suppose that a thief steals a car and he is captured. Should he just compensate the victim by the amount of money that the market says the car is worth (as is the case in, for instance, automobile accidents)? Or should the penalty be higher? It can be argued that the penalty should be higher than the market value so as to deter this type of transaction. But then the question becomes, why would we want to deter such transactions? One possible answer is that theft is not a regular market transaction in that it involves an externality that is not reflected in the market price of the object being stolen. One key difference between market exchanges and theft is that one of the parties in the theft does not participate voluntarily. But then again, this presumably also applies to automobile accidents. I think that this is an important question that has not yet been resolved in the literature of law and economics. See Klevorick (1985) for a survey.

¹⁰In fact this could be relaxed and the fee could be allowed to be concave as long as it is not too concave. The exact condition is $f'' > -(\beta - \omega) \cdot f' \cdot (1 + \pi) / [\pi \cdot c^{np}]$.

erroneously prosecuted and convicted). Since I am assuming that the fee people pay when they commit no crimes is zero (as $f(0)=0$) and, in addition, they do not suffer any disutility from being penalized, then whether innocent people are penalized or not is irrelevant (ie $c^P(t_i=0)=c^{NP}(t_i=0)$). Another way to think about the constant probability model is the following: every person faces a probability π of being investigated or searched by the police. If searched, the police find out how much crime that person has committed and, accordingly, he has to pay a fee. If it turns out that he did not commit any crimes, he pays nothing.

Individuals choose t_i so as to maximize utility (1) subject to (2), (3) and (4). The first-order condition entails the equalization of the marginal utility of t_i to zero. This condition can be rewritten as:

$$(5) \quad MB = (\beta - w/y) \cdot [1 - \lambda \cdot f(t_i)] = MC = \pi \cdot \lambda \cdot (c^{NP}/y) \cdot f'(t_i).$$

If we assume that the maximum possible fee is *all income* ($\lambda f(t_i) < 1$), then the term inside the squared brackets is positive. Since the right hand side of (5) is positive, people will devote positive amounts of effort to criminal activities only if $\beta > w/y$. In other words, only if the reward to committing crimes is higher than the reward of spending the same time in a legal activity will people commit crimes. This of course implies that only poor, low wage people will become criminals (rich people can earn more money by working).¹¹ The second order condition that ensures this is a

¹¹This does not mean that poor people are inherently worse in any sense. I have assumed that everybody has the same preferences towards crime and, therefore, everybody is equally good. The implication of the model comes from the opportunity set faced by both rich and poor. It is more profitable for the rich to be legal and for the poor to be criminal.

Of course I have assumed that the only reward for criminal behavior is the average level of income. It is entirely possible that rich people have access to a better, more rewarding set of criminal activities (white collar crime). If I amended the model to incorporate these factors, the implication would be that, given the size of the criminal reward a particular person faces, he would choose to devote zero time

maximum is

$$(6) \quad \partial^2 U / \partial t_i^2 = -(\beta - w/y) \cdot f'(t_i) \cdot (1 + \pi) - \pi \cdot (c^{np}/y) \cdot f''(t_i) < 0$$

In Figure 1 I plot the marginal benefit MB (which corresponds to the left-hand side of (5)) and marginal cost MC (which corresponds to the right-hand side of (5)) of criminal behavior. Because the fee is convex, the marginal cost is upward sloping. The marginal benefit is downward sloping. The optimal amount of criminal activity is determined by the crossing of MB and MC. If they cross at a point where t_i is between zero and one, the solution will be interior. If they cross to the right of $t_i=1$, individuals will become full-time criminals. If they cross to the left of $t_i=0$, individuals will devote all their time to legal activities.

The effects of growth on crime.

Imagine that the average income, the transfer received by potential criminals and the wage rate all increase in the same proportion. The first-order condition says that the amount of crime remains unchanged. In other words, in an economy where the fines, wages and transfers are fully indexed, the amount of crime is invariant to the level of income. The reason is that the rewards and costs of engaging in criminal behavior increase in the same proportion and, therefore, there is no additional incentive or disincentive to perform such activities.

Non-fully-indexed penalty systems, on the other hand, will tend to generate more crime as the economy grows since the rewards for committing crimes grow faster than the penalties. In terms of my analysis, this would correspond to a steady decline of λ holding everything else constant (I will analyze this case later

to illegal activities if the wage rate he can earn in legal activities is higher.

on). The model, therefore, has no direct prediction on the relation between the amount of crime and the level of income of the economy.

Increase in Income Inequality.

The process of economic development is sometimes not homogeneous across people: income inequality may increase or decrease as the economy develops. Some people argue that there is an inverse-U shape relation between income and inequality (Kuznets curve). We can analyze the effects of an increase in income inequality on the optimal amount of crime. In the present set-up this can be thought of as a reduction in the wage rate, w , holding constant the average level of income y , or a reduction in w/y . The MB schedule in Figure 1 shifts up while MC shifts down. The result is an increase in the optimal amount of crime. This can be also seen by applying the implicit function theorem to the first-order condition:

$$(7) \quad \partial t_i / \partial (w/y) = \frac{[1-\lambda \cdot f] + (1-t_i) \cdot f' \cdot \pi \lambda}{\partial^2 U / \partial t_i^2} < 0$$

The intuition is that an increase in income inequality reduces the benefits of working in the legal sector, while keeping the gains from crime constant. The obvious optimal reaction is an increase in crime. Hence, models that predict that economic growth is associated with larger income inequality will also predict an increase in disruptive activities. Ehrlich (1973) provides evidence supporting this proposition.

Better Enforcement of Laws.

Consider now an increase in the probability of conviction. This could be the result of higher investment in police protection or an improvement in the technology used by the police force. In terms of Figure 1, the MC line shifts upward while MB

remains unchanged. The total amount of crime goes down. The exact change is given by

$$(8) \quad \partial t_i / \partial \pi = \frac{(c^{np} / y) \cdot f' \cdot \lambda}{\partial U / \partial t_i^2} < 0.$$

Again the intuition is straightforward: a higher probability of being caught and convicted lowers the expected rewards of criminal activity and, therefore, lowers the number of crimes committed.

Larger Fees.

Imagine now that the authorities decide to increase the fees paid for every level of crime. This corresponds to an increase in λ in the model. The MB schedule shifts down and MC shifts up. The result is a reduction in the amount of crime. The quantitative change is given by

$$(9) \quad \partial t_i / \partial \lambda = \frac{(\beta - w / y) f + \pi \cdot c^{np} \cdot f' / y}{\partial^2 U / \partial t_i^2} < 0.$$

When penalties for being convicted are high, crime is low.

More Transfers and/or wage subsidies.

Finally, consider the effect of an increase in transfers (while maintaining average income constant). Because of the linear homogeneity of the fee with respect to income, the marginal benefit of committing crimes does not change. The marginal cost, on the other hand, increases as people who are convicted forego a larger

amount of income. The result is a reduction in crime:

$$(10) \quad \partial t_i / \partial (T/y) = \frac{\lambda \cdot f' \cdot \pi}{\partial^2 U / \partial t_i^2} < 0$$

Transfers in this model act just like fees since **they increase the (opportunity) cost of being penalized**: when convicted, people lose a fraction $\lambda f(t_i)$ of their income. Of course, the more they earn the more they lose if convicted. In other words, transfers provide an incentive to stay away from criminal activities by increasing the level of income outside jail. Hence, governments may want to use transfers as a mechanism to bribe people out of crime: **when transfers are high, crime does not pay**.

Note that this result depends on an increase in transfers relative to income. A certain amount of transfers protect the population against crime, given the amount of income. Income is the prize that criminals obtain by committing crimes. Holding constant the 'degree of protection' (transfers), an increase in the prize (income) induces people to commit more crimes.

Using cross-country data for 40 developed and developing economies, Tabellini (1992) finds that the level of transfers per unit of GDP is positively related to the pre-tax level of income inequality, even after he holds constant the initial level of income and the ratio of elderly to total population (both variables are significantly positively related to the level of transfers). He provides a political economy explanation for this finding. The theory outlined in this paper, however, is also consistent with these correlations: income inequality leads to high levels of crime and, therefore, to the need for public welfare protection.

A natural question to ask is why and when would governments go to the trouble of establishing a tax/transfer system instead of just increasing penalties, given

that transfers act just like penalties or fees? To answer this question we must bear in mind that there are limits to the fees that governments can impose on people. In particular, people cannot pay more than everything they own.¹² Suppose that the penalty system is such that the fees paid if caught being a full-time criminal ($t_1=1$) are **everything**.¹³ Consider that group of people (**desperate people**) whose wage rate relative to the average is so low that, despite these enormous fees, they decide to become full time criminals (so they pay everything if caught). An increase in fees will not induce these desperate people out of criminal behavior because they will already lose everything if convicted. Hence, once people are in such a desperate situation, fees are irrelevant in the sense that higher fees will not decrease criminal behavior. Transfers, on the other hand, will still work as an incentive device to reduce crime because they are not a direct cost but rather an **opportunity cost** to committing crimes: by increasing the amount of income people receive if they stay out of jail, transfers increase the size of 'everything' to criminals. Hence, they still increase the penalty and, therefore, they still reduce the optimal amount of crime.

Note that, in this model where there is no leisure choice, **wage subsidies** work in much the same way that transfers do. A wage subsidy would increase w relative to y . We already established that an increase in w/y reduces crime. Thus, like transfers, wage subsidies work as a crime-reduction device. Note that, also like

¹²Here is where the assumption that governments cannot impose non-monetary penalties like death or cutting off people's ears becomes relevant. Presumably the value of lives and ears in terms of income is large enough so that crime can be deterred with the use of these non-monetary penalties only. Countries that have access to these types of drastic penalties (and some Muslim countries do) will not need to use transfers to reduce disruptive behavior. In this paper I will not try to explain why governments do not impose such big non-monetary penalties for seemingly small crimes.

¹³People cannot lose exactly everything when they go to jail: the government must provide some level of consumption while in jail. If this was not the case, prisoners would starve to death. This would represent a non-monetary penalty which I assumed was not allowed in this economy. This sentence should therefore say that they lose 'almost' everything.

transfers, what matters is the wage rate relative to the average level of income in the economy. As we saw above, if wages and income increase in the same proportion (along with transfers and fees) the total amount of crime will remain unchanged.

An additional point is that, when people find it optimal to commit crimes under a certain economic environment, it is likely that they will still find it optimal to commit crimes after serving time in jail unless the economic environment has changed. Transfer programs and public subsidies may be a way to change this adverse economic environment.¹⁴

Making the Probability of Conviction a Function of Crime.

Up to now I have assumed that every person was investigated by the police with the same probability π . This probability was independent of the amount of crime committed. It is natural to assume that the probability of capture and conviction is increasing in the amount of crime a person decides to commit: in the real world, the probability of non-criminals being arrested by mistake is not zero but it is surely smaller than the probability faced by true criminals. As my grandmother used to say: 'when you are playing with fire,...you are going to get burnt'. Following my grandmother's infinite wisdom, I now will assume that π is an increasing function of t_i with $\pi'(t_i) > 0$, $\pi''(t_i) < 0$ and $\pi(0) = 0$. Individuals still maximize (1) subject to (2), (3), and (4), taking into account that their actions will

¹⁴This assumes that people don't learn anything new in jail. It could be the case that criminals did not really know what jail was all about and that an initial period of incarceration shows them how terrible it is. This would increase the perceived penalty and, therefore, reduce the amount of crime in the future. One argument against this is that a lot of criminals come from families and neighborhoods where crimes and criminals are abundant. Hence, it is likely that these people have a pretty good idea of what it is to be in jail so their propensity to commit crimes will not change after having been in jail once before. (see Sah (1991) or evidence on this type of social osmosis).

affect the probability of being caught:

$$(5)' \quad (\beta - (w/y)) \cdot [1 - \lambda \cdot f(t_i)] - \pi \cdot \lambda \cdot [(w/y) \cdot (1 - t_i) + \beta t_i + (T/y)] f'(t_i) - \\ - \pi'(t_i) \cdot \ln(1 - \lambda f(t_i)) \cdot (1 - \lambda f(t_i)) \cdot [(w/y) \cdot (1 - t_i) + \beta t_i + (T/y)] = 0.$$

The first two terms are the same as in (5). They represent what would be optimal if the probability of capture was unaffected by the choice of t_i . The third term reflects the marginal losses in utility due to the increase in the probability of capture when people decide to devote one more unit of time to illegal activities. Note that this first-order condition is still invariant to the level of income if the wage rate and the transfer system are fully indexed (that is, if w/y and T/y are constant). Hence, growth that preserves income inequality still does not have an effect on the level of crime. Using the implicit function theorem, we can see that crime is still increasing in income inequality and decreasing in the size of the penalties. The effect of transfers on crime, on the other hand, is now the following:

$$(10)' \quad \partial t_i / \partial (T/y) = \frac{\pi \cdot \lambda \cdot f' - \pi' \cdot (1 - \lambda f) \cdot \ln(1 - \lambda f)}{\partial^2 U / \partial t_i^2} < 0,$$

where $\partial^2 U / \partial t_i^2$ is negative according to the second-order conditions (which are satisfied if $\pi'' > 0$ or if π'' is not too negative). The numerator of (10)' is positive: the first term is the product of three positive numbers. The second term is the negative of a product of positive numbers times $\ln(1 - \lambda f)$. Since both λ and $f(t_i)$ are positive fractions, the number inside the logarithm is less than one and, therefore, the logarithm is negative. Therefore, T/y still acts as a crime-reducing device.

The main lesson is that if we allow the probability of capture and conviction to be an increasing function of the amount of crime committed, the relevant features of

the model do not change. In particular, transfers are still an opportunity cost of being penalized and, therefore, they act as a crime-preventing device.

A Model With Leisure Choice.

The simple model used up to now treats wage subsidies and transfers in a very symmetric way. The reason is that agents were not allowed to choose the amount of leisure optimally. One could argue that if the choice of leisure is allowed, then a transfer induces people to want to buy more leisure. Of course they do so by reducing the time spent in the activity with the lowest reward: legal work. Wage subsidies (which you can collect only if you work), have an offsetting substitution effect as the relative reward of legal work. Transfers that are not linked to work, however, do not have the substitution effect while they still have the perverse wealth effect. To investigate whether this perverse effect is possible in my model, let me amend the utility function so as to incorporate a preference for leisure:

$$(1)' U = \pi \cdot \ln(c^P) + (1-\pi) \cdot \ln(c^{np}) + \psi \cdot \pi \cdot \ln(l^P) + \psi \cdot (1-\pi) \cdot \ln(l^{np}),$$

where ψ is some discount rate on leisure, l^P is the amount of leisure the agent enjoys if penalized and l^{np} is the leisure the agent enjoys if not penalized. The time spent working is $(1-t_i-1)$, where t_i is still the time devoted to crime (because total time available is still normalized to 1). As in the previous section we define c^{np} and c^P as follows:

$$(2)' \quad c^{np} = w \cdot (1-t_i-1) + \beta \cdot y \cdot t_i + T$$

$$(4)' \quad c^P = [1-\lambda f(t_i)] \cdot c^{np}.$$

I will assume that part of the penalty for criminal behavior is in terms of lost

utility. If we denote the amount of leisure enjoyed when not penalized by l (so $l^{np}=l$), the leisure enjoyed when penalized is:¹⁵

$$(11) \quad l^P = [1 - \lambda f(t_i)] \cdot l.$$

Agents choose l and t_i so as to maximize utility subject to the constraints (2)', (3)', and (11). The first-order conditions entail

$$(12) \quad (\beta - w/y) \cdot [1 - \lambda f(t_i)] - (1 + \psi) \cdot \lambda \tau \cdot f'(t_i) \cdot [(w/y) \cdot (1 - t_i - l^*) + \beta t_i + T/y] = 0,$$

where l^* is the optimum amount of leisure given by

$$(14) \quad l^* = \frac{\psi}{1 + \psi} \cdot [(1 - t_i) + \frac{\beta t_i + (T/y)}{w/y}].$$

The derivative of l^* with respect to the transfer per unit of income is positive

$$(15) \quad \partial l^* / \partial (T/y) = \frac{\psi}{1 + \psi} \cdot (y/w) > 0.$$

Other things being equal, more transfers lead people to enjoy more leisure. Using (15), we can now calculate the effect of an increase in transfers per unit of income

¹⁵We could also assume that the fraction of income lost if convicted is different from the fraction of time lost if convicted. The reader can check that the key results remain the same.

on crime

$$(16) \quad \frac{\partial t_i}{\partial(T/y)} = \frac{(1+\psi)\lambda\pi f' [1-(w/y) \cdot \partial l^* / \partial(T/y)]}{\partial^2 u / \partial t_i^2} =$$

$$= \lambda \cdot \pi \cdot f' / [d^2 u] < 0,$$

where $d^2 u \equiv -(1+\pi)\lambda f'(\beta-(w/y)) - (1+\psi)\lambda \pi f'' c^{np}/y < 0$. The first term inside the squared brackets reflects the negative effect of transfers on crime that we outlined in previous sections. The second term inside the brackets $(w/y) \cdot \partial l^* / \partial(T/y)$ reflects the perverse wealth effect that transfers have on the consumption of leisure and, as a result, on crime. Under my particular specification, the overall effect of transfers on crime is unambiguously negative. That is, the perverse wealth effect never dominates. This result does not depend on the log utility specification (the overall effect with a utility function of the form $c^{1-\theta}/(1-\theta)$ and $l^{1-\theta}/(1-\theta)$ yields:

$$\frac{\partial t_i}{\partial(T/y)} = \frac{[\{\pi(1-\lambda f)^{1-\theta} + (1-\pi)\}w]^{1/\theta} \cdot \pi \lambda f'}{[(\pi(1-\lambda f)^{1-\theta} + (1-\pi)w)^{1/\theta} + (\pi\psi)^{1/\theta} \cdot w] \partial^2 u / \partial t_i^2},$$

which is still negative.

The effect of wage subsidies on crime, on the other hand, involves no potentially perverse effects. The reason is that, unlike transfers, wage subsidies have a negative effect on leisure: public transfers in the form of wages, increase the reward to legal activities. The substitution effect induces an increase in work effort and a reduction in crime and leisure. The wealth effect involve an increase in leisure and a reduction in crime and work. The overall effect is a reduction in

leisure. The overall effect on crime is given by

$$(17) \quad \partial t_i / \partial (w/y) = \frac{(1-\lambda f)(w/y) + (1+\psi)\lambda \pi f' [(1-t_i-1^*) - \{\partial 1^* / \partial (w/y)\}(w/y)]}{\partial^2 u / \partial t_i^2} < 0,$$

where $\partial 1^* / \partial (w/y) = -\frac{\psi}{1+\psi} [(\beta t_i + T/y) / (w/y)^2] < 0$. Note that all the terms in the numerator of (17) are positive while the denominator is negative. Hence, there is no perverse wealth effect from wage subsidies.¹⁶

The lesson from this section is that, even though we could think that transfers that are not linked to work may have a perverse effect on criminal behavior due to a wealth effect on leisure, the overall effect is still negative. However, the quantitative effects of wage subsidies on crime are likely to be much larger than those of transfers. The main result is still that public welfare should have a negative impact on the amount of time people devote to criminal activities and that the relevant variable is the total spending on public welfare as a ratio to the average income of the economy (which is, in turn, related to the average prize of criminal behavior).

¹⁶In a general equilibrium model, wage subsidies may have another perverse effect on crime, as they tend to generate unemployment. Note that this is not the case for transfers.

Public Welfare, Taxes, and Growth.

In the previous sections I considered the partial equilibrium effects of aggregate public welfare policies on the criminal behavior of people. The natural question to ask is, given that the government can reduce crime by increasing the size of the public welfare system, why doesn't it get rid of all crime by having an enormous public welfare program?. The answer is, of course, that transfers and subsidies need to be financed by raising taxes. Taxes, in turn, may distort private choices for savings and investment which, in turn, affect the consumption path. The government, therefore, will have to balance the distortionary effects of the implicit 'taxes' imposed by criminals with those of the explicit taxes imposed by the government itself. In this section I use a simple model of growth in order to analyze these issues.

Agents maximize a utility function of the form

$$(18) \quad \int_0^{\infty} e^{-\rho t} \frac{c^{1-\theta}-1}{1-\theta} dt,$$

where c is the average consumption of the population. There are two ways to think about (18). First we could think that the representative agent does not care about the utility of criminals. Under this interpretation, c is the average consumption of the non-criminal population. Alternatively, we would think in terms of the veil of ignorance of Harsanyi and Rawls where, ex-ante, people do not know whether they will end up being criminals or not. If we assume that, ex-ante, all agents are identical, the choice variable c could be interpreted as the level of consumption of the representative or average agent, and (18) then represents his utility.

I will imagine that, as a result of criminal and disruptive activities, some aggregate output is lost. Since most crimes entail just a transfer from victim to

criminal (at least this is true for most property crimes) one could think that no aggregate output is lost as a result. There are several reasons, however, why output losses may exist. First, society may not care about the happiness of criminals. If this is the case, any resources that end up in their hands should be considered social losses. Second, victims of crime may be emotionally and physically disrupted. The consequence of such disruption will be a reduction in the victim's ability to perform his job at pre-crime levels. Crime, therefore, lowers labor productivity. Third, private individuals may devote effort, time, and resources to protect themselves against crime. This is a social waste much in the same manner as rent-seeking activities that use up some output for no particularly useful purpose. Fourth, some output may be simply destroyed as a result of criminal activities: at the very least, robbers are careless and they break precious pieces of china when they enter somebody's house. The worst possible scenario is of course the loss of life as a consequence of a simple robbery (of course widespread fires and destruction are intermediate cases of loss of GDP).

A fraction $1-\varphi(\cdot)$ of income is lost and a fraction $\varphi(\cdot)$ is still available after crime. We can also think of $\varphi(\cdot)$ as the instantaneous probability of maintaining one's property rights on output. According to the analysis above, this fraction or probability will be an increasing function of the overall level of police and legal protection, an increasing function of the size of the penalties for conviction, and a decreasing function of income inequality. Most importantly, it will be an increasing function of the total amount of aggregate transfers or public welfare, TR, per unit of average income.¹⁷ Since I assume that the population is constant, I can normalize the stock of people to one so average and aggregate income coincide. In order to

¹⁷We should think of TR as including not only transfers but also wage subsidies and other kind of public welfare. As we showed in previous sections, all of them affect crime negatively. In the rest of the paper, I use the terms transfers and public welfare interchangeably.

concentrate on the effects of transfers on growth, let me assume that the fraction φ is solely a function of TR/Y^{ac} , where Y^{ac} is 'after-crime' national income.¹⁸ In particular, I neglect police protection and public investment in property rights and law enforcement, despite the fact that these are expenditures relevant to criminal activities. Hence, I assume

$$(19) \varphi(\cdot) = \varphi(TR/Y^{ac}), \quad \varphi'(\cdot) > 0, \quad \varphi''(\cdot) < 0, \quad \varphi(0) \geq 0, \quad \varphi(1) \leq 1, \quad \varphi(0) < \varphi'(0),$$

where the assumption on the last inequality is made so as to ensure that the problem of crime is important enough to warrant public intervention.

Under this specification, redistributational transfers and public welfare resemble productive public goods subject to congestion: the amount of income people get to keep after crime depends on the level of public welfare relative to the size of criminal threat. This threat, in turn, depends on the prize that criminals get if they decide to commit crimes, which is proportional to national income. When a person increases his economic activity, he raises the economy's average level of income and, with it, it congests the protective power of public welfare (Thompson (1974) and Barro and Sala-i-Martin (1992) interpret national defense spending along similar lines).

I will imagine that the production function is linear in the capital stock

$$(20) \quad y^{pc} = A \cdot k$$

where y^{pc} is per capita "pre-crime" income and k per capita capital. The linearity

¹⁸Alternatively, it could be assumed that $\varphi(\cdot)$ is a function of TR per unit of pre-crime income. This alternative specification does not change any of the substantive results.

of the production function is not essential to my analysis but it enables me to get closed form solutions for growth rates (I could use a neoclassical production function and the growth effects of different policies would be temporary and analytically intractable; the direction of the growth effects along the transitional path would, however, be the same).

The government collects revenue from a constant tax rate on after-crime income τ (I assume that illegal income does not pay taxes) and always runs a balanced budget. All components of public spending other than transfers are excluded from the present analysis. All public revenue is therefore spent on public welfare. The government budget constraint is:

$$(21) \quad TR = \tau \cdot \varphi(TR/Y^{ac}) \cdot AK$$

where K is the aggregate capital stock and $Y^{ac} = \varphi(\cdot)AK$ is after-crime aggregate income. After-tax output is devoted to either consumption or investment. The constraint faced by the individual is, therefore

$$(22) \quad \dot{k} = (1-\tau) \cdot \varphi(TR/Y^{ac}) \cdot A \cdot k - c - \delta \cdot k$$

where $k_0 > 0$ is given, τ is the tax rate on 'after-crime' output, and δ is the constant rate of capital depreciation. Individuals maximize (18) subject to (22). Since all agents are small relative to the aggregate, they all think that their actions do not affect the behavior of the government. Hence, when they optimize they take

$\varphi(\text{TR}/Y^{\text{ac}})$ and τ as given.¹⁹ The first-order conditions are

$$(23) \quad \gamma_c \equiv \dot{c}/c = [(1-\tau) \cdot \varphi(\text{TR}/Y^{\text{ac}}) \cdot A \cdot - \rho - \delta] / \theta$$

$$\lim_{t \rightarrow \infty} p \cdot k = 0$$

where p is the shadow price associated with the constraint (22). Note that, since τ is constant, the government budget constraint says that TR/Y^{ac} is also constant so transfers and output grow at the same rate. It follows that consumption grows at a constant rate at all points in time. From the budget constraint (22), it can be seen that in the steady state, consumption, physical capital and, therefore, output grow at the same rate $\gamma_c = \gamma_k = \gamma_y = \gamma_{\text{TR}} \equiv \gamma$. The transversality conditions imply that physical capital grows at that same rate at all points in time. Hence, the model displays no transitional dynamics as all variables grow at the same constant rate all the time.

We can use (23) to find the growth rate of the economy²⁰

$$(24) \quad \gamma = [(1-\tau) \cdot A \cdot \varphi(\tau) - \rho - \delta] / \theta.$$

The size of the public welfare program has two effects on the growth rate γ : on the one hand higher taxes reduce growth as they distort investment decisions (this is the term $(1-\tau)$ in (24)); on the other hand they increase growth as they reduce the

¹⁹I assume that individuals, who own the firms, produce output at home. The results would be the same if there were competitive markets for goods and capital.

²⁰If we assume that $\varphi()$ is a function of TR/Y rather than TR/Y , the growth rate is not a function of $\varphi(\tau)$ but, instead, a function of $\eta(\tau)$ with $\eta'(\tau)$. Where $\eta()$ can be derived as follows: define $\varphi_2()$ as the function that satisfies the public budget constraint $\varphi_2(\text{TR}/\text{AK}) \equiv \tau$ where $\varphi_2'() > 0$ (this follows from the assumptions $\varphi'' < 0$ and $\varphi(0) \geq 0$). Invert it and plug in $\varphi(\text{TR}/\text{AK})$ to get the growth rate as a function of τ only where $\eta(\tau) = \varphi(\varphi_2^{-1}(\tau))$. Since both $\varphi()$ and $\varphi_2()$ are monotonically increasing, $\eta'(\tau) > 0$.

amount of crime and disruption in the economy (this is the term $\varphi(\tau)$ in (24)). For high levels of τ (large governments) the first, detrimental effect dominates. For low levels of τ the second, beneficial effect dominates because $\varphi(0) < \varphi'(0)$ (see (19)). In words, if the crime problem when there is no public welfare is important enough, then an increase in the size of such public programs will increase the growth rate of the economy. If this condition does not hold (so the crime problem is not important), then it could be the case that $\partial\gamma/\partial\tau < 0$ for all τ so the optimal size of the government is $\tau^* = 0$. Under these circumstances, there is a size of the government τ^* at which the two effects exactly cancel out and growth reaches its maximum. The rate τ^* is given by the following implicit function

$$(25) \quad \partial\gamma/\partial\tau = -\varphi(\tau^*) + (1-\tau^*) \cdot \varphi'(\tau^*) = 0.$$

The maximization of growth is not always equivalent to the maximization of the utility of the representative agent. This is true, however, when $\varphi(\cdot)$ takes a Cobb Douglas form.

Superiority of Income Taxes.

It is interesting to compare the outcome of this market economy with that of a planner. Given an arbitrary size of the government, $s = TR/[AK \cdot \varphi(TR/AK)]$, the planner chooses a path of consumption and capital so as to maximize the utility of the representative consumer. The resulting growth rate is

$$(26) \quad \gamma_{PL} = [(1-s) \cdot A \cdot \varphi(s) - \rho - \delta]/\theta$$

where the effect of s on the rate of growth is given by the following expression

$$(27) \quad \partial \gamma_{PL} / \partial s = -\varphi(s^*) + (1-s^*) \cdot \varphi'(s^*) = 0.$$

Note that if a government using income taxes chooses τ so as to maximize growth ($\tau = \tau^* = s^*$), then the social optimum will be replicated. In other words, if the size of the government is optimal, the proportional income tax is Pareto efficient. It is interesting to see that a shift from income tax to lump-sum tax lowers utility but it increases growth. The growth rate under lump-sum taxes is given by:

$$(28) \quad \gamma_{LS} = [A\varphi(s) - \rho - \delta] / \theta,$$

Given the size of the government, the growth rate corresponding to lump-sum taxes is always larger than the one the planner would choose. That is, if taxes are lump sum there is overinvestment and excessive growth. The intuition for this result is that, when an individual producer decides to increase capital by one unit, he increases the average output of the economy. This in turn induces criminals to increase their criminal effort since the rewards for crime have increased. In other words, investors congest the protective role of transfers without really taking this into consideration when making investment decisions. Therefore, they tend to overinvest and overcrowd transfers. A lump-sum tax will not do anything to solve this congestion effect. An income tax, on the other hand, acts as a fee for the use of transfers as a crime-preventing device: it internalizes the externality and deters people from investing too much. Thus, from a social point of view, an income tax is superior to a lump-sum tax.

Empirical Evidence Using Cross-Country Data.

Let me finish by providing some empirical evidence in favor of the theory developed in this paper. The key difference between my model and all other models of transfers is that mine suggests that, because they are 'productive', transfers should affect growth positively. We can check these predictions using cross-country data: IMF's Government Financial Statistics provide data on government transfers for a sample of 75 countries going back to 1970.²¹ I use the ratio of transfers to GDP, along with the growth rates in real per-capita income between 1970 and 1985 from the Summers and Heston data set. I also use the 1970 ratio of public consumption and public investment to GDP and the savings rate constructed by Barro (1991).

The model predicts that, in a cross-country regression, the growth rate of the economy should be positively related to the size of the transfer program, once the size of the government (which in the model is reflected by τ) and proxies for the preference parameters are held constant. This prediction is checked in the first column of Table 1. I proxy the preferences towards savings with the initial (1970) savings rate and the size of the government with the 1970 ratio of total spending to GDP. Initial income is included so as to allow for the possibility of transitional dynamics (recall that I assumed an Ak technology for simplicity but if, instead, the technology is neoclassical, the growth implications of the model are the same for a transitional period. This transition is reflected in the initial level of income). The coefficient is negative and significant ($-.0128$ s.e.=0043), which reflects the importance of the transitional process. The key coefficients, however, are the ones on the size of

²¹The GFS transfer variable also includes old-age pensions. In Sala-i-Martin (1992), I show that old-age pensions should ALSO be regarded as productive as they induce unproductive, old people out of their jobs. Hence, I am not too worried about the fact that this may be too broad a measure of transfers. Nevertheless, I think it would be interesting to distinguish empirically which one of the two components of total transfers dominates the results. For most poor countries of this sample, however, separate data on redistributive and intergenerational transfers is not available.

the government, $-.1117$ (s.e.=.0370) and the transfer to GDP ratio $.1092$ (s.e.=.0509). As expected, holding the overall size of the government, transfers are positively related to per-capita growth. I repeat the experiment in column 2, using the log of TR/Y instead of the level. The overall results are the same (the coefficient on $\log(TR/Y)$ is $.0050$ (s.e.=.0018)).

In the model outlined above, I neglected public investment, public consumption and other forms of public spending. Following Barro (1990), we could easily include such variables in the model. As in Barro (1990), the additional predictions would be that productive spending (such as public investment) should be positively related to growth, while non-productive spending (such as public consumption) should affect growth negatively. In column 3, I break the total size of the government into three components: public consumption (GC/Y), public investment (GI/Y) and transfers (TR/Y). Public consumption enters negatively ($-.1285$, s.e.=.0475) and public investment is insignificant ($-.2278$, s.e.=.1728). The only variable that is positively related to growth is transfers, with a coefficient of $.1108$ (s.e.=.0522).

These results suggest that, contrary to the predictions of ALL other theories, transfers are positively related to growth rates for a large cross-section of countries.

Conclusions.

In this paper I presented a model that explains the existence of redistributive transfers. I showed that such transfers are a mechanism to buy poor people out of disruptive activities such as crime, revolutions or week-long inner city riots. I argued that public welfare is likely to have some effect on crime, especially among those segments of the population that are so poor that the losses of going to jail are very small relative to the potential gains from criminal behavior. I also argued that, in aggregate production functions, transfers and other forms of welfare look like productive public inputs subject to congestion which increase the productivity of private capital and, therefore, increase the growth rate of the economy. I then derived the growth-maximizing size of the public welfare system. I showed that, as a result of transfers being 'subject to congestion', an income tax system was superior to a lump-sum tax system. The reason was that income taxes act as user fees on the use of welfare as a protective device. Finally, I provided international evidence in favor of the model: contrary to the predictions of all other theories of transfers, the data suggest that other things equal, countries that have larger transfers programs tend to grow faster.²²

²²Using panel data for a sample of 23 OECD countries, Cashin (1992) also finds a positive partial relation between the size of the transfer program and the rate of growth.

Table 1: Growth and Transfers

	(1)	(2)	(3)
	GR7085	GR7085	GR7085
Constant	.0007 (.0099)	.0251 (.0138)	.0002 (.0111)
ln(GDP70)	-.0128 (.0043)	-.0133 (.0037)	-.0147 (.0049)
τ	-.1117 (.0370)	-.1093 (.0356)	
Savings Rate	.2006 (.0357)	.1997 (.0373)	.2168 (.0515)
TR/Y	.1092 (.0509)		.1108 (.0522)
log(TR/Y)		.0050 (.0018)	
GC/Y			-.1285 (.0475)
GI/Y			-.2278 (.1728)
R ²	.35	.37	.35
s.e.	.0183	.0179	.0182

Notes: Standard Errors are in parenthesis. All regressions have been estimated using White's heteroscedasticity-consistent covariance matrix. The data for GDP and growth rates is from Summers and Heston (1988). GR7085 is the annualized growth rate of per capita GDP. ln(GDP70) is the logarithm of the 1970 per capita GDP. τ is a measure of the 1970 ratio of total government spending to GDP and is taken from Barro (1991). The savings rate is the 1970 ratio of total investment to GDP. R/Y is the average of the ratio of social security transfers to GDP for the period 1970-1985. GC/Y and GI/Y is the ratio of total government consumption (excluding defense and education) and total investment to GDP for 1970. They are taken from Barro (1991). Sample size: 75 countries.

References.

- Alesina, A. and D. Rodrik (1991), "Redistributive Politics and Economic Growth", NBER WP# 3668.
- Barro, R. J. (1990), "Government Spending in a Simple Model of Endogenous Growth", *Journal of Political Economy*, 95, S103-125, October.
- Barro, R. J., (1991), "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics*, May .
- Barro, R. J. and X. Sala-i-Martin, (1992), "Public Finance in Models of Economic Growth", mimeo Harvard University, forthcoming *Review of Economic Studies*.
- Becker, G. S., (1968), "Crime and Punishment: An Economic Approach" *Journal of Political Economy*, April.
- Blanchard, O. J. and S. Fischer (1991), "Lectures in Macroeconomics", MIT press, Cambridge, MA.
- Cashin, P., (1992), "The Effect of Public Investment, Transfers, and Taxes on Economic Growth: Theory and Evidence for 23 Developed Countries", Mimeo Yale University, May.
- Chamley, C., (1981), "The Welfare Cost of Capital Income Taxation in a Growing Economy", *Journal of Political Economy*, 89, 468-496.
- Ehrlich, I., (1973), "Participation in Illegal Activities", *Journal of Political Economy*, May.
- Klevorick, A. K., (1985), "Legal Theory and the Economic Analysis of Torts and Crimes", *Columbia Law Review*, 85, June, 905-920.
- Lucas, R. E. Jr. (1990), "Supply-Side Economics: An Analytical Review", *Oxford Papers*, 42, 293-316.
- Persson, T. and G. Tabellini, (1991), "Is Inequality Harmful for Growth?", mimeo Berkeley.
- Sala-i-Martin, X., (1992), "Transfers" mimeo Yale University, March.
- Sah, R. K., (1991), "Social Osmosis and Patterns of Crime", *Journal of Political Economy*, 99, 6, December, 1272-1295.
- Stigler, G., (1970), "The Optimal Enforcement of Laws", *Journal of Political Economy*, 78, 2, March/April.

Summers, R. and A. Heston, (1988), "A New Set of International Comparisons of Real Product and Price Levels: estimates for 130 countries", *Review of Income and Wealth* 34, March, 1–25.

Tabellini, G. (1992), "A Positive Theory of Social Security", mimeo IGIER, Milano.

Thompson, E. A., (1974), "Taxation and National Defense", *Journal of Political Economy*, 82, July/August, 755–782.

Waldfogel, J. (1992), "The Effects of Criminal Conviction on Income and the Trust Reposed in the Workmen", mimeo Yale University, 1992.

Figure 1

