Education and Crime over the Life Cycle*

Giulio Fella[†]

Giovanni Gallipoli[‡]

Abstract

We develop an overlapping-generation, life-cycle model with endogenous education and crime choices. Education and crime depend on different dimensions of heterogeneity. We apply the model to property crime and calibrate it to U.S. data. We compare two policies: subsidizing high school completion and increasing the length of prison sentences. We find that targeting crime reductions through increases in high school graduation rates entails large efficiency and welfare gains. These gains are absent if the same crime reduction is achieved by increasing the length of sentences. The cost-effectiveness of high school subsidies increases significantly if they are targeted at the wealth poor. We find that general equilibrium effects explain half of the reduction in crime from subsidizing high school and are non-negligible even for interventions targeted at low levels of wealth. Crucially, the effect of small equilibrium price changes is magnified by their interaction with the underlying individual heterogeneity.

JEL classification: H52

Keywords: Crime, education, equilibrium, policy.

*We thank Sami Berlinski, Francesco Caselli, Marco Cozzi, Lance Lochner, Iourii Manovskii, Antonio Merlo, Costas Meghir, Nicola Pavoni, Victor Rios-Rull, Jean-Marc Robin, Kjetil Storesletten and Gianluca Violante for comments and discussions. We are grateful to Lance Lochner for making available the NLSY data on crime and to Robert Valletta for his data on age-dependent unemployment duration and rates. Giulio Fella would like to thank the University of Pennsylvania for the hospitality while working on this paper.

[†]Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS, UK. E-mail: `g.fella@qmul.ac.uk`

[‡]Department of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, BC V6T 1Z1, Canada. E-mail: `gallipol@interchange.ubc.ca`

1 Introduction

Crime is a hot issue on the U.S. policy agenda. Despite its significant fall in the Nineties its cost to the taxpayer has soared. The prison population has doubled over the same period and now stands at over two millions of inmates. The yearly cost of keeping a person in jail exceeds 20,000 dollars.[1] The dramatic increase in the U.S. prison population over the past twenty years has prompted a shift of interest, among both academics and policymakers, from tougher sentencing to other forms of intervention.[2]

This paper develops a heterogeneous-agent, equilibrium life-cycle model incorporating both education and criminal choices. Its goal is to provide a framework within which to compare the effectiveness and, importantly, the welfare implications of alternative policies which directly or indirectly impact on crime. The model emphasizes the crucial role of heterogeneity. Agents differ in: 1) innate, observed ability, 2) innate crime propensity[3] and 3) initial wealth. Agents self-select themselves into education on the basis of these differences and, upon entering the labour market, decide whether to engage in criminal activity on a period-by-period basis.

We apply the model to the study of property crime which - unlike, for example, violent crime - is most likely to be driven by economic considerations. We calibrate the model to data for the US economy. We verify that the model is consistent with a number of untargeted features of the data by running identical estimation procedures on simulated data from our benchmark economy as well as comparable data from the NLSY79.

We use our model to evaluate two alternative policies: a subsidy towards high school completion and an increase in the prison sentence which generates the same change in the crime rate as the subsidy policy. The size of the subsidy is 8 per cent of average labour earnings per year of schooling and is chosen to coincide with the corresponding value in a well-known, small scale program: the Quantum Opportunity Program. The program

[1]The figure comes from Pastore and Maguire (1995) and is at the lower end of available estimates. Section 3 reports alternative estimates.

[2]See, e.g., the survey by Donohue and Siegelman (2004).

[3]Ability is a set of characteristics that directly affects earnings, whereas crime propensity residually captures in a single index the range of unmodelled individual, family and social characteristics that contribute to shaping individual attitudes towards crime (e.g. gender, broken families, neighbourhood quality). Merlo and Wolpin (2009) provide evidence supporting the importance of residual heterogeneity in accounting for criminal behaviour.

provided extra support and high school graduation incentives aimed at children from a disadvantaged background.[4]

We first consider the effect of making the subsidy available to everybody completing high school. The policy increases the equilbrium share of high school graduates in the population by one percentage point and reduces the crime rate by 4.5 per cent relative to the benchmark. Even more importantly, the subsidy implies significant efficiency and welfare gains. Efficiency, as measured by the steady state flow of aggregate consumption, increases by 0.6 per cent. Welfare, as measured by the permanent consumption equivalent of ex ante, expected lifetime utility, increases by nearly 2 per cent. The marginal impact of the policy decreases as the size of the subsidy is increased to up to three times its value in the Quantum Opportunities Program. A subsidy equal to 24 per cent of average labour income reduces the crime rate by 8 per cent, relative to the benchmark, and raises efficiency by 1.5 per cent and welfare by about 4 per cent.

Compared to an unconditional high school subsidy, an increase in the prison term that induces the same fall in the crime rate generates no efficiency gain and only negligible welfare gains. Intuitively, the efficiency gains of the subsidy come from its effect on the education composition of the labour force. No such effect is present in the case of the prison term. Concerning welfare, the only effect of the increase in the prison term is to reduce the transitory income risk associated with being the victim of a crime. On the other hand, by increasing the relative price of labour for high school dropouts and weakening the link between wealth at birth and selection into education, the subsidy provides insurance against the (ex-ante) uncertain ability and initial wealth draws. Since the ability shock is permanent and the initial wealth shock has a persistent effect through the education choice, the welfare benefits are large as they cumulate over the whole lifetime.

Conducting the same subsidy experiment in partial equilibrium reveals that the general equilibrium increase in the relative price of high school dropouts, although small, plays an important role in terms of crime reduction. Even small changes in prices have significant effects through their interaction with the underlying heterogeneity. The crime fall associated with a subsidy equal to 8 per cent of average labour income is only half as

[4]See Hahn, Leavitt, and Aaron (1994) and Taggart (1995) for a discussion of the program and its effects.

large as in general equilibrium.

Comparing the effect of an unconditional subsidy to that of a subsidy conditional on wealth reveals that the bulk of the efficiency and welfare gains of the unconditional subsidy are due to the response of the poorest individuals. A means-tested subsidy restricted to people in the lowest thirtieth percentile of the wealth distribution achieves nearly three quarters of the efficiency and welfare gains of the unconditional subsidy and between half and three quarters of the reduction in crime.

Conducting the same means-tested experiment, but randomizing over a treatment and a control group allows us to assess the effectiveness of the intervention on different groups within a target population and to explicitly control for selection when evaluating the program impact. We find that treated individuals experience a 5 per cent drop in crime over the life cycle relative to an appropriately defined control group.

The model is in the tradition of economic models of crime which goes back to Becker's (1968) seminal work. It builds upon the original contribution of Imrohoroglu, Merlo, and Rupert (2004) who are the first to construct a calibrated, structural, general equilibrium model of rational crime choice.[5] Their model is extremely successful in accounting for the evolution of the U.S. property crime rate over the last twenty-five years on the basis of changes in wage inequality, employment opportunities, age and education distributions and expected punishment. The focus of their analysis is positive. With the aim of accounting for changes in the crime rate, they take the education distribution as exogenous and let it vary according to its evolution in the data over the relevant period. We extend Imrohoroglu, Merlo, and Rupert's (2004) framework by endogenizing investment in education and the marginal returns to education.

In a related contribution Imrohoroglu, Merlo, and Rupert (2000) endogenize police expenditure and the degree of redistribution by adding a political economy dimension within an equilibrium model of crime. The model is able to account for several features of U.S. data, such as the positive correlation between police expenditure and redistribution and the lack of correlation between redistribution and crime. Cozzi (2006) also uses a

[5]For reasons similar to those highlighted in Heckman, Lochner, and Taber (1998), the analysis of alternative policies to tackle crime benefits from the use of a dynamic equilibrium framework. The case for the use of models allowing for equilibrium effects in policy analysis has been recently argued by various authors in different fields. See, among others, Abraham (2001), Lee (2001), Lee and Wolpin (2004), Cunha, Heckman, and Navarro (2004) and Gallipoli, Meghir, and Violante (2009).

calibrated equilibrium model to investigate the interaction between drug addiction, drug crime and property crime. He uses the model to quantify the effects of two policies: the compulsory treatment of drug offenders and drug legalization. He finds that both policies have a significant impact on property crime.

There is an extensive body of empirical literature testing the main prediction of the rational theory of crime that both market returns and the expected punishment are significant determinants of criminal choices. The effect of market returns upon crime is well established: see for example Grogger (1998), and Freeman (1999) who surveys earlier empirical studies. More recent work includes Gould, Weinberg, and Mustard (2002), Machin and Meghir (2004), Raphael and Ludwig (2003) and Tobacman (2009). Concerning the effect of the expected punishment, Levitt (1997) and Levitt (1996) find sizeable elasticities of crime rates respectively to expenditure on police and the length of the expected prison term.

The existence of a relationship between crime and education is documented by Lochner and Moretti (2004). They find that high school graduation significantly reduces participation in both property and violent crime. On the basis of their estimated elasticities and assuming that general equilibrium effects are negligible, they calculate that a 1 per cent increase in high school graduation would imply large savings in terms of reduced costs from crime. We find that general equilibrium effects significantly reinforce the effect of increased high school graduation even if we assume an elasticity of substitution between different types of human capital at the top end of the range of available estimates. In fact, equilibrium effects account for half of the crime reduction associated with an increase in high school graduation in our model.

Donohue and Siegelman (2004) assess the cost-effectiveness of alternative policies aimed at tackling crime, including social policies. Their cost-benefit analysis, though, relies on elasticities from existing empirical studies.

The structure of the paper is the following. Section 2 introduces the model. Section 3 discusses the model parameterization. Section 4 simulates the model, assesses its performance and studies the effect of alternative policies. Section 5 concludes.

2 The Economic model

The model has an overlapping generation structure. Time is discrete.

2.1 Demographics and preferences

The economy is populated by a continuum of individuals. At each date a new cohort of unit mass starts life. We denote by $j \in J = \{0, 1, ..., \bar{j}\}$ the age of an individual and by λ_j the conditional probability of surviving from age j to $j+1$. The unconditional probability of surviving up to age j is $\Lambda_j = \Pi_{s=0}^{j-1} \lambda_s$.[6]

Individual preferences over consumption of the unique final good are time-separable. They are captured by the felicity function $u(c)$ and the discount factor $0 < \beta < 1$.

2.2 Production technology

Firms are identical and use human and physical capital to produce output according to the production function

$$Q(H, K) = H^{1-\eta} K^{\eta}, \qquad 0 < \eta < 1, \tag{1}$$

where H and K denote, respectively, the aggregate stocks of human and physical capital. The human capital stock H is the aggregate

$$H = [\sigma_0 H_{e_0}^{\rho} + \sigma_1 H_{e_1}^{\rho} + (1 - \sigma_0 - \sigma_1) H_{e_2}^{\rho}]^{\frac{1}{\rho}}. \tag{2}$$

of the stocks of human capital with education $\{e_0, e_1, e_2\}$. Workers with the same level of educational attainment are perfect substitutes.

Physical capital depreciates at the exogenous rate δ.

2.3 Government

The government administers a pay-as-you-go pension system, the criminal justice system, spends on wasteful public expenditure and transfers and collects taxes. It balances the

[6]By the law of large numbers, Λ_j is also the mass of agents of age j in the population.

budget at all times.

The government pays a pension benefit b to each pensioner and bears a total cost m for each convicted criminal. It also pays a subsidy[7] sub_e to a student with wealth a for each period spent studying towards a degree e. Both pension benefits and student subsidies are tax-exempt while labour and capital income are taxed at the proportional rates t_l and t_k respectively.

In the model benchmark, once the transfers and the criminal justice systems have been financed, any excess tax revenue is spent on non-valued public expenditure G.

2.4 Market arrangements

Markets for the four factors of production and the unique final good are competitive. There are no state-contingent markets to insure against income risk, but workers can self-insure by borrowing and saving into a risk-free asset subject to an exogenous borrowing limit $\underline{a}$.

We denote by r the riskless interest rate and by w_e the wage per efficiency unit of labour of a worker with education e. Both prices are net of the respective taxes.

Absent markets to insure against mortality risk, the economy features involuntary bequests. The distribution of involuntary bequests determines the distribution of wealth at birth according to the following mechanism. An inheritance tax equal to $\min\{a, a^*\}$ is levied on any bequest amounting to a. If $a < 0$, the tax is actually a transfer which is paid to the creditors of the deceased.[8] The quantity a^* ensures that all bequeathed debts are repaid in full and solves

$$\int_{\underline{a}}^{a^*} a d\Gamma(a) + a^* \int_{a^*}^{\infty} d\Gamma(a) = 0, \tag{3}$$

where $\Gamma(a)$ is the distribution of bequeathed wealth.

The intuition behind equation (3) is the following. The sum on the left hand side is the total net revenue from the inheritance tax. This has to be zero, since the only

[7]For expositional convenience, we do not allow for means-tested subsidies in the description of the model, though we do allow for them in the numerical experiments in sections 4.4 and 4.5.

[8]This ensures that the asset is indeed risk free. In the absence of this pooling of the mortality risk, risk free lending would imply a (natural) borrowing constraint at zero.

purpose of the tax is to repay the creditors of people dying in debt. The first integral is the net revenue from all bequests $a \leq a^*$. The contribution to net revenue is actually negative, it is a claim by creditors, if $a < 0$. The second integral is the revenue from bequests $a > a^*$. If the borrowing limit $\underline{a}$ is negative, a^* has to be positive to ensure that all creditors are paid. If instead the borrowing limit $\underline{a}$ is non-negative, no individual dies in debt and no inheritance tax is required to ensure that all creditors are paid. In such a case, the solution to equation (3) is $a^* = 0$.

The tax implies that any pre-tax bequest a maps into a post-tax bequest $\max\{0, a - a^*\}$. The distribution of wealth a at birth $\Gamma^0(a)$ is assumed to coincide with the distribution of post-tax bequests; namely if one denotes by $\Gamma(a)$ the distribution of pre-tax bequests it is

$$\Gamma^0(a) = \begin{cases} \Gamma(a + a^*) & \text{for } a \geq 0 \\ 0 & \text{for } a < 0. \end{cases} \tag{4}$$

If $\underline{a}$ is non-negative, hence a^* equals zero, the distribution of wealth at birth coincides with the distribution of (pre-tax) bequests. If, instead, $\underline{a}$ is negative, the wealth distribution at birth is a translation of the distribution of pre-tax bequests censored at zero.

The borrowing limit $\underline{a}$ is actually negative in our parameterization for the reasons discussed in Section 3.2.

2.5 Life cycle

Individuals are born with a common education level e_0, but are heterogeneous along the following dimensions. They differ in their initial stock of financial wealth a, in their innate ability θ and in their utility cost χ of engaging in crime. The three shocks are independently drawn.

To ease notation we stack the two shocks (θ, χ) in the vector $\xi \in \mathbb{R}^2$. We denote by $G(\xi)$ the cumulative density function for ξ.[9]

[9] Given that the shock ξ is permanent, we subsume it in the value functions in the description of the model that follows, in order to streamline notation.

2.5.1 Education

Individuals choose their education level $e \in \{e_0, e_1, e_2\}$. At the start of life, they make a discrete choice between entering the labour market as a high school dropout ($e = e_0$) or studying towards a high school degree ($e = e_1$). The latter choice entails attending high school for j_{e_1} years. At age j_{e_1} a high school graduate chooses between entering the labour market or studying towards a college degree ($e = e_2$). College lasts $j_{e_2} - j_{e_1}$ years. Upon completing college at age j_{e_2}, individuals enter the labour force.

Studying towards a degree e entails a one-off utility cost $\psi_e(\theta)$. The heterogeneity in the disutility of education captures, in reduced form, heterogeneity in environmental and other unmodelled factors that contribute to the cross-sectional variability in educational achievement. For example, Heckman, Lochner, and Todd (2006) discuss the importance of (heterogeneous) psychic costs to reconcile observed enrollment rates and market returns to education.

The out-of-pocket cost of studying toward degree e equals the fee f_e minus a subsidy sub_e for each year attended. For tractability, we do not allow any flexibility in the timing of education choices and assume full commitment to the completion of an ongoing education cycle[10]. For the same reason, we do not allow individuals to work or engage in crime while in education.

Let $V_j^s(e, a_j)$ and $V_j^n(e, a_j)$ denote respectively the value of being or not in education for an individual of age j, (omitted) type ξ, completed education e and financial wealth a_j.

The value function of a newborn with initial wealth a_0 satisfies

$$V^0(e_0, a_0) = \max\{V_0^s(e_0, a_0) + \psi_{e_1}(\theta), V_0^n(e_0, a_0)\}, \tag{5}$$

as the agent chooses optimally between attending high school and entering the labour market.

The value function of a student with degree e studing towards degree e' satisfies the

[10] Since there is no accrual of information while in school, a time-inconsistency problem arises only if students can drop out without repaying the education subsidy received.

Bellman equation

$$V_j^s(e, a_j) = \max_{c_j, a_{j+1}} u(c_j) + \beta\lambda_j V_{j+1}^s(s, e, a_{j+1}) \tag{6}$$

subject to

$$a_{j+1} = a_j(1+r) + f_{e'} - sub_{e'} - c_j, \quad a_{j+1} \geq \underline{a}, \tag{7}$$

if $j \neq j_{e_1} - 1$, and

$$V_j^s(e, a_j) = \max_{c_j, a_{j+1}} u(c_j) + \beta\lambda_j \max\{V_{j+1}^s(e', a_{j+1}) + \psi_{e_2}(\theta), V_{j+1}^n(e', a_{j+1})\}, \tag{8}$$

subject to (7), if $j = j_{e_1} - 1$. The continuation value for a student in the last year of high school is the result of the optimal choice, next period, between continuing education or entering the labour market. We denote the optimal choice, at age j_e, of whether to study or not towards degree e' by[11]

$$i_{j_e}^s = \begin{cases} 1 & \text{if } V_{j_e}^s(e, a_j) + \psi_e(\theta) \geq V_{j_e}^n(e, a_j) \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

2.5.2 Work and crime

Upon entering the labour market, individuals supply their labour endowment inelastically until the compulsory retirement age j_r. Labour supply l_j is subject to independently-distributed employment shocks. It equals $0 < \underline{l}_j < 1$ with age-dependent probability $\pi_j^u > 0$ and one with the complementary probability.[12]

The labour efficiency of an individual is a function

$$h_j(\theta, e) = \exp(\gamma(e)\theta + \varpi_j(e)) \tag{10}$$

with an education-specific ability gradient $\gamma(e)$ and age component $\varpi_j(e)$. Only a completed education cycle affects effective labour supply.

In every period, individuals can engage in criminal activity regardless of their em-

[11] Since the highest attainable degree is college, the student's problem is subject to the condition $V_{j_{e_2}}^s(e_2, a_j) = V_{j_{e_2}}^n(e_2, a_j)$.

[12] This way of modelling unemployment shocks allows us to contain the computational burden by using a time interval of a year in our calibration while still being able to accommodate a shorter unemployment duration. The latter is around one quarter for the average U.S. worker.

ployment status. Criminal activity amounts to theft and has a utility cost χ. Though modelled as a utility cost, the permanent shock χ is meant to capture, in reduced form, heterogeneity in initial conditions that are uncorrelated with measured ability and family wealth and affect an individual's propensity to engage in crime. In a sense it plays the same role as an individual fixed effect and residually captures all forms of heterogeneity such as gender, neighbourhood, race, broken families that we do not explicitly model. For this reason, we refer to it as the crime fixed effect in what follows. [13]

An individual of age j chooses the number of thefts τ_j he commits from the discrete set $\{0, 1, ..., \bar{\tau}\}$ and is apprehended and sent to prison with probability $\tau_j \pi_p$. To streamline exposition we assume that the length of the prison term is one period.[14]

For a victim, a theft involves losing a fraction α of post-tax labour income. Namely, it is a multiplicative shock $v \in \{0, \alpha\}$ to labour income with $\pi_v = Pr\{v = \alpha\}$ the probability of being a victim. For simplicity we assume criminals cannot target their victims and that each theft yields a fraction α of the average post-tax labour income.[15]

The consumption/saving choice takes place at the end of the period, after all uncertainty, including that concerning apprehension, has been resolved. A convicted criminal receives no labour income - and cannot be robbed - but keeps her assets and the proceeds from her last crime. While in jail, a criminal cannot access her wealth and consumes the exogenous amount $\bar{c}$.

It follows that an individual's dynamic budget identity satisfies

$$a_{j+1} = a_j(1+r) + \tau_j \alpha \bar{y} + (1 - i^p)[(1 - v)\, w_e h_j\, (\theta, e) - c_j], \tag{11}$$

where i^p is a random variable which takes value one if $\tau_j > 0$ and the individual is convicted, and zero otherwise. Individuals who are not in prison in the current period earn labour income but can be robbed of a share v of it. The term $\tau_j \alpha \bar{y}$ is the illegal return from committing n crimes.

[13] The importance of allowing for individual heterogeneity in subjective returns from criminal activity was first emphasized by Flinn (1986). Our modelling approach is consistent with Merlo and Wolpin (2009) who find that permanent, unobserved heterogeneity in initial conditions plays a crucial role in explaining the crime and schooling experience of young black males.

[14] In fact, we do not impose the restriction in the quantitative part.

[15] The modelling choice implies that only workers can be the object of a crime. It has the advantage that, unlike wealth, labour income is always positive.

Let the superscript p (for prison) index a convicted criminal and let $\mathbb{E}_{l_j}$ and $\mathbb{E}_v$ denote the expectation operators with respect to the probability distribution of labour supply shocks l_j and the victimization shock v. The value function of a worker before observing her current labour supply realization is

$$V_j^n(e, a_j) = \mathbb{E}_{l_j}\left[\max_{\tau_j}\ \chi\mathbb{I}_{\tau_j} + \tau_j\pi_p V_j^p(e, a_j) + (1 - \tau_j\pi_p)\mathbb{E}_v[\max_{c_j, a_{j+1}} u(c_j) + \beta\lambda_j V_{j+1}^n(e, a_{j+1})]\right], \tag{12}$$

where $\mathbb{I}_{\tau_j}$ is an indicator function equal to one if $\tau_j > 0$ and zero otherwise. Upon observing her labour supply shock l_j, an individual chooses how many crimes to commit. If $\tau_j > 0$, he bears a utility cost χ and is apprehended with probability $\tau_j\pi_p$, in which case her value function equals

$$V_j^p(e, a_j) = u(\bar{c}) + \beta\lambda_j V_{j+1}^n(e, a_{j+1}). \tag{13}$$

Any individual not in jail is subject to the random shock v associated with being robbed and chooses consumption after observing the shock realization.

2.5.3 Retirement

Individuals retire from the labour force at age j_r and receive a lump-sum, yearly pension b until the end of their life. Their value function satisfies

$$V_j^n(e, a_j) = \max_{c_j, a_{j+1}} u(c_j) + \beta\lambda_j V_{j+1}^n(e, a_{j+1}), \tag{14}$$

$$\text{s.t. } a_{j+1} = a_j(1 + r) + b - c_j, \tag{15}$$

$$a_{j+1} \geq \underline{a}, \quad a'_{\bar{j}} \geq 0,$$

where the terminal condition $a_{\bar{j}+1} \geq 0$ rules out insolvency in the last period of life.

2.6 Stationary Equilibrium

Let $z_j^x \in Z_j^x$ denote the state implicit in the recursive representation of the problem for an individual of age j and type x, where $x = s$ for a student and $x = n$ otherwise.

For a given set of government policies $\{b, G, sub_e, t_l, t_k\}$ and apprehension probability π_p, a *stationary recursive equilibrium* is a collection of (i) policy functions for consumption and saving $\{c_j(z_j^x), a_{j+1}(z_j^x)\}$, education $\{i_{j_{e_0}}^s(z_j^s), i_{j_{e_1}}^s(z_j^s)\}$ and crime $\{\tau_j^n(z)\}$; (ii) value functions $\{V_j^s(z_j^s), V_j^n(z_j^n), V^{n,p}(z_j^n)\}$; decision rules $\{K, H_{e_0}, H_{e_1}, H_{e_2}\}$ for firms; (iv) prices $\{r, w_{e_0}, w_{e_1}, w_{e_2}\}$; (v) a victimization rate π_v; (vi) an average labour income $\bar{y}$; (vii) time-invariant measures $\{\mu_j^s, \mu_j^n\}$ and $\Gamma(a)$; (viii) a wealth level a^* that satisfy the following conditions.

1. Given prices$\{r, w_{e_0}, w_{e_1}, w_{e_2}\}$:
    - the decision rules $\{c_j(z_j^x), a_{j+1}(z_j^x)\}$ and the value functions $\{V_j^s(z_j^s), V_j^n(z_j^n), V_j^{n,p}(z_j^n)\}$ solve equations (6)-(7) for students, equations (11) and respectively (12) and (13) for non-students in and out of prison of age $j < j_r$, and (14)-(15) for agents of age $j \geq j_r$;
    - the education decisions $\{i_{j_{e_0}}^s(z_j^s), i_{j_{e_1}}^s(z_j^s)\}$ solve equation (9);
    - the crime decision $\tau_j(z_j^n)$ solves equation (12).
2. Given prices$\{r, w_{e_0}, w_{e_1}, w_{e_2}\}$, input demands $\{K, H_{e_0}, H_{e_1}, H_{e_2}\}$ maximize profits for the representative firm
$$r = (1 - t_k)(F_K - \delta)$$
and
$$w_e = (1 - t_l)F_{H_e}, \text{ for } e \in \{e_0, e_1, e_2\}.$$
3. The asset market clears
$$K = \sum_{j,x} \int_{Z_j^x} a_{j+1}(z_j^x) d\mu_j^x$$
.
4. The labour markets for each educational level clear[16]
$$H_{e_i} = \sum_{j<j_r} \int_{\{z_j^n : e=e_i\}} l_j h_j(\theta, e)(1 - \pi_p \tau(z_j^n)) d\mu_j^n, \text{ for } i \in \{0, 1, 2\}.$$
where the supply of labour on the right hand side of the above equation is made up only of individuals out of jail. These are a fraction $(1 - \pi_p \tau(z_j^n))$ of workers in their

[16]By Walras law, market clearing on all factor markets ensures that the goods market clears.

age group, in stationary equilibrium.

5. The government budget is balanced

$$G + E + PRIS + PENS = \frac{t_k}{1-t_k} rK + \frac{t_l}{1-t_l} \sum_e w_e H_e$$

Total government outlays on the left hand side of the above equation are the sum of exogenous wasteful expenditure G, education subsidies $E = \sum_{j,i} \int_{\{z_j^n : e=e_i\}} sub_{e_i} \, d\mu_j^s$, aggregate prison expenditure[17] $PRIS = \sum_j \int_{Z_j^n} m \, \pi_p \tau(z_j^n)) d\mu_j^n$ and aggregate pension expenditure $PENS = \sum_{j \geq j_r} \int_{Z_j^n} p \, d\mu_j^n$.

6. The victimization rate coincides with the crime rate

$$\pi_v = \left(\sum_{j<j_r} \int_{Z_j^n} (1 - \pi_p \tau(z_j^n)) d\mu_j^n \right)^{-1} \sum_{j<j_r} \int_{Z_j^n} \tau_j(z_j^n) d\mu_j^n,$$

and equals the total number of crimes divided by the total number of workers out of jail.

7. The average disposable labor income satisfies

$$\bar{y} = \left(\sum_{j<j_r} \int_{Z_j^n} (1 - \pi_p \tau(z_j^n)) d\mu_j^n \right)^{-1} \sum_e w_e H_e.$$

8. The distribution of pre-tax bequests satisfies

$$\Gamma(a) = \sum_{j,x} \int_{\{z_j^x : a_{j+1}(z_j^x) \leq a\}} (1 - \lambda_j) \mu_j^x$$

9. The vector of measures $\mu = \{\mu_0^s, ..., \mu_{\bar{j}}^s; \mu_0^n, ..., \mu_{\bar{j}}^n\}$ is the fixed point of $\mu(Z) = Q(Z, \mu)$ where Z is the generic subset of the Borel sigma algebra $\mathfrak{B}_Z$ defined over the state space $\mathbb{Z} = \prod_{j,x} Z_j^x$, the Cartesian product of all Z_j^x. The mapping $Q(Z, \mu)$ is the transition function associated with the individual decisions, the law of motion for the shocks $\{\xi, v, l_j\}$ and the survival probabilities $\{\lambda_j\}$.

10. The cutoff wealth level a^* solves (3).

[17] In stationary equilibrium, the number of convicted felons in each age group equals a fraction $\pi_p \tau(z_j^n))$ of the corresponding number of workers.

3 Parameterization

We select parameter values using a combination of external sources and calibration, matching simulated model moments with corresponding moments in U.S. data for the early Eighties. The chosen period is dictated by data availability, as we use the 1979 wave of the NLSY to calibrate the distribution of innate ability and its relationship to earnings. Parameter values are reported in tables 1 to 6.

3.1 Parameters from external sources

Demographics. Each period represents one full year. An individual is born at age sixteen as a high school dropout. High school takes two years to complete. When eighteen a high school graduate can study for four additional years in order to acquire a college degree. In any case, agents can work only until age sixty-five. The age range in the model is the same as the age range we use in our NLSY sample. The maximum possible age in the model is ninety-five. The survival probabilities prior to age ninety-five are taken from NCHS (1997).

Unemployment shocks. The probability π_j^u of experiencing a fall in labour supply is chosen to equal the unemployment rate for the relevant age group. The associated fall in labour supply $1 - \underline{l}_j$ is chosen to match the duration of unemployment in the data. For example, the unemployment rate for the average U.S. worker is 6.5 per cent with an unemployment duration of one quarter. This would correspond to $\pi^u = 0.065$ and $\underline{l} = 0.75$ in our model. The age-specific duration and unemployment rates are taken from[18] Valletta (2005) and Daly, Jackson, and Valletta (2007) and are constant within the age-brackets [16,19], [20,24], [25,34], [35,54] and [55,64].

Production technology. The capital depreciation rate δ and the share of capital income η are set respectively to 0.065 and 0.35 (see Cooley, 1995). The parameters σ_0, σ_1 and ρ of the human capital aggregator in equation (2) rely on the estimates by Gallipoli, Meghir, and Violante (2009), based on CPS and PSID data. They estimate a value for ρ between 0.36 and 0.68, corresponding to an elasticity of substitution between 1.6 and

[18] We are grateful to Robert Valletta for kindly providing his dataset.

3.1.[19] In the simulations we set $\rho = 0.68$, which implies relatively high substitutability and conservative general equilibrium effects. The share parameters are calibrated using an average of the values estimated for the period 1979-1981.[20] One point deserves some discussion: the marginal products of human capital for different education groups depend on the aggregate human capital supplies, which are determined in equilibrium by the optimal choices of agents. Given the estimated technology parameters, the utility costs of education are chosen to match the shares of workers in each educational categories, as discussed below.

Distribution of innate ability. So as to be able to simulate our model, we need to approximate the distribution of unobserved ability affecting wages and education choices. To this purpose we follow Gallipoli, Meghir, and Violante (2009) and use the cross-sectional distribution of recorded AFQT test scores as reported in the NLSY79. This data set has the advantage of providing a measure of cognitive skills for a representative sample of an entire cohort in the U.S. population, as well as direct wage measures over time. We can therefore link measured 'ability' and wages within different education groups. In Figure 5 we report the measured AFQT distribution for the whole cross-sectional sample of the NLSY79. For computational simplicity, the distribution is approximated by grouping all agents in 5 bins (quintiles) containing equal proportions of the total population. The range of each bin is different, as its extremes correspond to successive quintiles of the empirical ability distribution. Ability is assumed to be uniformly distributed within each bin.

Wages. Individual wages are determined as the marginal product of human capital in a given education group times the efficiency units of an agent. Consistent with equation (10) in our model, efficiency units for each education group e are a function of age $\xi_j(e)$ and permanent ability θ with education-specific, ability gradient $\varpi_j(e)$. In order to quantify the relative effect of these two components, we use estimates from Gallipoli,

[19] Using a skilled/unskilled classification Katz and Murphy estimate the elasticity of substitution in production to be 1.41, while Heckman, Lochner, and Taber (1998) report a favorite estimate of the elasticity of substitution between skilled and unskilled equal to 1.44. Card and Lemieux (2001) find that the elasticity of substitution between different age groups is large but finite (around 5) while the elasticity of substitution between college and high school workers is about 2.5.

[20] See Table 1 for estimates of the share parameters for several years; details of the estimation procedure can be found in Gallipoli, Meghir, and Violante (2009) and are available on demand.

Meghir, and Violante (2009). Estimates for the ability gradient are based on NLSY79 data. Using information on the highest grade completed, individuals are allocated to three education groups: high school drop-outs (HSD), high school graduates (HSG) and college graduates (CG). Table 5 reports the ability gradient for different education groups, under the assumed specification. The age component of individual efficiency are estimated from PSID data, which provide a longer working life span and are presented in Table 5.

Government. The government levies proportional taxes on capital and labour. Following Domeij and Heathcote (2003), we set $t_l = 0.27$ and $t_k = 0.4$. For simplicity, the pension is assumed to be a constant lump sum for all agents, regardless of their education and previous earnings. The pension replacement rate is set to 16 per cent of average pre-tax labor earnings as in Heathcote, Storesletten, and Violante (2008).

Direct Cost of Education. The direct cost of college education is chosen to match the value of the average tuition costs net of the average grant as a proportion of average pre-tax earning for the academic year 1980-81. The figure for the average tuition cost is the average "Tuition and required fee" over all 4-year institution from the National Center for Education Statistics. The corresponding average grant size is from Lewis (1989). The resulting net cost equals 9.7 per cent of average pre-tax earnings. As for the yearly cost of attending high school, we set it to be just 1 per cent of average pre-tax earnings, in order to account for expenses incurred for study material and other practicalities. There does not seem to be not much information on such costs.

3.2 Calibrated parameters

Preferences. We choose a CRRA specification for the function $u(c)$. We set the coefficient of relative risk aversion to 1.5, in the middle of the range of available estimates (see, for example, the survey by Attanasio, 1999).

The size of the wealth stock is a crucial determinant of consumers' ability to smooth income shocks. For this reason, it is important for the model to generate a wealth income ratio comparable to that in the data. We set the discount factor β to match a wealth-income ratio of 2.7. Wolff (2000) reports that this is the ratio between total net worth and total pre-tax income for the average household, excluding the top 5 per cent percentile, in the 1983 Survey of Consumer Finances. The reason for excluding the top 5 per cent of households is that the PSID and NLSY, from which most of our estimates are derived, significantly undersample the top of the wealth distribution.[21] The implied value for the discount factor is 0.958.

Borrowing Limit. The exogenous borrowing limit $\bar{a}$ is calibrated to match the share of workers (all agents excluding students) with zero or negative wealth. Wolff (2000) provides an estimate of 15% for this share in 1983, which implies a negative borrowing limit of about 26 per cent of average pre-tax labour earnings in our benchmark. Note that in the absence of the mechanism for pooling mortality risk discussed in Section 2.4, the natural borrowing constraint would be zero. Riskless borrowing and lending would require that no individual die in debt with probability one. Our calibration and modelling choice implies a borrowing constraint which is looser than the natural one but, importantly, is consistent with agents' aggregate borrowing patterns in the data.

Education Enrolment Rates. Education rates are matched both in the aggregate and by ability groups. The distinction is important because the same aggregate shares are consistent with different distributions of ability over education and, therefore, many different relative marginal returns between different types of labor. Moreover, alternative policies are likely to alter the ability composition in each education group. Therefore it is essential that the benchmark reproduces the distribution of ability types over education outcomes.

Education shares for each of the five ability bins in which we have partitioned the range of AFQT scores are measured using NLSY79 data. However, the aggregation of education shares based on the NLSY79 does not match the aggregate education distribution of workers in the U.S. economy, as measured by the CPS for 1980. [22] For aggregate consistency we proportionally adjust the NLSY rates so that they aggregate to the average CPS education rates observed between 1977 and 1983[23]. This grossing-up is such that

[21]For example, Kaplan and Violante (2009) make a similar argument.

[22]The reason for this discrepancy is that the NLSY79 refers only to one cohort of the U.S. population, whereas the CPS gives a snapshot of the education distribution of workers at all ages.

[23]The aggregate education distribution for workers between 1977 and 1983 was close to the average for the period 1967-2001, which is the sample period used in Gallipoli, Meghir, and Violante (2009) to estimate the wage equations and the production technology.

enrolment in each education group is rescaled equally across all ability bins, so that the relative ability composition of each bin in unchanged relative to that in the NLSY79. The education shares by ability and their aggregate value are reported in Table 6.

Finally, the ability-specific quasi-linear utility terms $\psi_{j_e}(\theta)$ are calibrated to match the grossed-up joint distribution over education and ability.

Crime parameters. Our definition of property crime is the same as in the data. It is any crime which qualifies as burglary, larceny or motor-vehicle theft. Our measures of income from crime, probability of incarceration and length of sentence are averages of the corresponding measures in each of these categories with weights equal to the relative frequency of each of them.[24]

We set α, the share of average labour income obtained from one crime, to match its counterpart in the data. From Table 3.76 in Pastore and Maguire (1982), the average loss for a victim of one property crime was 728 dollars in 1980. Average labour income in the same year was 12,400 dollars, based on our CPS sample. The resulting value for α is 5.87 per cent.

The probability of being convicted for committing one crime is the product of the probability that a crime is cleared by arrest and the suspect indicted (clearance rate), the probability of facing trial conditional on being indicted and the probability of being sentenced to prison conditional on being put on trial. These probabilities were respectively 16.8, 80 and 44 per cent in 1980[25], which imply $\pi_p = 5.7$ per cent.

The sentence length is set to 19 months, which is the average time served in 1983, the closest survey year, by prisoners convicted for property crime offenses released from state prisons.[26] Following Cozzi (2006), we model the fractional prison term as a lottery between a prison term of one and two years with respective probabilities that imply an expected prison sentence of 19 months.[27]

We assume that the innate disutility of engaging in crime has a uniform distribution

[24]The weights change with the variable under consideration; e.g. they are the relative frequency of each category of property crime in the case of income from crime, but the relative number of prisoners charged with each category of crime in the case of sentence length.

[25]Respectively from tables 4.19, 5.19 and 5.20 of Pastore and Maguire (1981).

[26]From Table 6.31 in Pastore and Maguire (1986).

[27]This implies that an individual committing crime in the last year of her working life spends the first retirement year in prison with positive probability. We assume that her pension is paid into her bank account, though he cannot access it until released.

over the interval $[a, b]$. We set a, b and the exogenous consumption level in prison $\bar{c}$ to match the property crime victimization rate and the shares of recidivists and high school dropouts among apprehended criminals. These were respectively 5.6, 61 and 52.7 per cent in 1980.[28] The chosen values are $a = -3.55$, $b = .17$, and $\bar{c} = 0.38$ of average labour income.

The expenditure per convict in the model is set to 67 per cent of average labour income. This corresponds to the average (across jails and state prisons) cost per prisoner in 1980.[29]

3.3 Computation

The model is solved numerically using a generalized version of the endogenous grid method developed in Fella (2009). The method extends the original idea of Carroll (2006) to environments with non-convex choice sets.[30] It relies on the standard endogenous grid method to locate local maxima and uses a comparison method to select the global maximum among them. The method provides a robust way to solve our problem, which entails a sequence of discrete choices over the whole life-cycle, and affords very large gains in computational time. To the best of our knowledge this is the first application of the method to non-convex problems.

4 Numerical Analysis

This section discusses the benchmark equilibrium and presents the results of our policy experiments. Section 4.1 describes the equilibrium for the benchmark economy in some detail. It also provides evidence that the model captures the basic forces determining property crime both at the individual and aggregate level, by comparing estimates of age- and education-dependent probabilities of crime participation from the simulated

[28]The first figure is from Table 3.53 in Pastore and Maguire (1982). The other two are, respectively, from "Examining Recidivism" (NCJ-96501), Bureau of Justice Statistics Special Report 1985, and "Profile of State Prison Inmates" (NCJ-58257), 1979.

[29]The figures for the average cost per jail and state prisoner are respectively from "Correctional Populations in the United States" (NCJ-156241), 1993 and "Federal and State Prisons. Inmate Populations, Costs and Projection Models" (B-272244), General Accounting Office, 1996.

[30]Barillas and Fernandez-Villaverde (2007) extend the endogenous grid method to perform value function iteration in models with more than one control variable, but with a convex choice set.

and NLSY79 data. The remaining sections describe the policy experiments. In all the experiments, the proportional tax rate on labour income adjusts to balance the government budget.

4.1 The Benchmark Economy

Selection into education. Both wealth and ability are important determinants of selection into education in the benchmark economy. Table 5 reports the average wealth, as a fraction of average labour income, for students and workers at the time of the high school and college choices. The difference in average wealth between students and workers in the same ability bin is particularly pronounced at the time of the high school choice. In particular, those individuals in the three highest ability bins that drop out of high school are extremely wealth poor.

Interestingly, at the age of college choice labour market entrants in the first two ability bins have higher wealth than their counterparts that enrol in college. This can be understood by noticing that selection takes place both on ability *and* wealth even within a given ability bin. Furthermore, returns to education, net of foregone wages while in school, are increasing in ability. The higher ability the lower must be the wealth of an individual at the margin between enrolling or not in a given schooling level. Less able students are richer at the margin. Selection into high school explains the negative selection along the wealth dimension at the time of college choice within the first two ability bins. Individuals who will continue to college are relatively more able, but relatively poorer (already at the time of high school choice) than their counterparts that will stop at high school.

For individuals in the top ability bins instead, returns to schooling are so high that at the margin there is effectively no trade-off between wealth and ability: among high school students wealth is basically uncorrelated with ability. At the time of the schooling choice selection takes place mainly along the wealth dimension. It is poorer individuals who choose not to study.

Evaluating the model performance. Our benchmark calibration has realistic implications for the composition of the criminal population, replicating a variety of facts which are not directly targeted.

First, the proportion of criminals who were unemployed at the time of conviction in the model is around 28 per cent. According to Table 6.31 in Maguire and Pastore (1983) 30 per cent of the state prison population in 1979 (the closest survey year to our calibration year) was unemployed at the time of incarceration. The fact that a majority of people engaged in crime are employed is also consistent with the evidence in Grogger (1998).

Second, we compare the age profiles of crime participation in the model and in the data. We use data from the 1980 wave of the NLSY79 survey. These data contain the same information on AFQT that we have used to calibrate our model and, more importantly, being individual-level data, they contain specific information on participation in property crime.[31]

The 1980 wave of the NLSY79 asked respondents, males between 15 and 23 of age at the time, a number of questions concerning participation in property and other types of crime. The measure of property crime we use is whether or not the respondent reported having engaged in shoplifting, or stealing something worth $50 or more, from someone/somewhere other than a store. The counterpart in the model data is whether an individual commits a positive number of crimes in a given period.

Figure 5 reports the average male participation rates for high school dropouts and high school graduates in the model and NLSY data. The NLSY dataset at our disposal covers only ages 18 to 23. For this reason, we cannot include standard errors for the participation rates in the NLSY data for ages 16 and 17. The point estimates are taken from Figure 1, chart (C) in Lochner (2004). In the model there are no high school graduates younger than 18.

Since the model is calibrated to the whole population and not just males the data need to be rescaled in order to be comparable to those from the NLSY. We calculate the rescaling factor in the following way. If one denotes by ϕ the ratio between the female and male participation rates, the aggregate participation rate equals the male rate times the correction factor $\mu = .5 \times (1 + \phi)$. We approximate ϕ by the ratio between female

[31] The same dataset has been used by Lochner (2004) and Lochner and Moretti (2004) to study the effect of education on crime. The dataset we use has been kindly provided by Lance Lochner. A large subset of the data is available on Enrico Moretti's webpage at http://www.econ.berkeley.edu/ moretti/data.html

and male jail inmates in 1982, the closest survey year. The ratio is 7 per cent according to Table 6.18 in Pastore and Maguire (1983).[32] We assume ϕ is constant across ages and rescale the model participation rates - the number of criminals divided by the size of the potential criminal population- by dividing them by the correction factor μ.

The black squares in Figure 5 are the average participation rates in the NLSY. The dashed lines trace the 95 per cent confidence bounds. The red triangles are the model counterparts. Reassuringly, the simulated and actual participation rates are not statistically different at the 95 per cent confidence level for all ages but 19.[33] As for college graduates, their participation rate is close to zero in the model. The corresponding number is about 5 per cent and not statistically different from zero in the NLSY data.

Third, we run identical OLS regressions on both the simulated and NLSY data for individuals aged 18-23.[34] The dependent variable is a dummy equal to one if the individual is engaged in property crime. We compare the estimated coefficients on a dummy for high school completion and the AFQT percentile across these regressions. Table 5 reports the estimated coefficients multiplied by 100. Columns (1) and (2) report results for the simulated data while (3) to (5) refer to the NLSY data. Columns (1) and (3) control only for age. The coefficients on the AFQT score is not economically significant, though statistically different from zero in column (1). Reassuringly, one cannot reject the null hypothesis that the coefficients from the model and NLSY data are equal. This is especially important for the coefficient on the high school graduation dummy, which is both economically and statistically significant. High school graduation is associated with an approximate 6 per cent reduction in the probability of engaging in property crime both in the model and NLSY data. This result indicates that the model generates a realistic relationship between education status and criminal behaviour and is robust to a set of specification changes. For example, column (2), (4) and (5) introduce controls

[32] According to Table 6.20 the ratio is even lower, 4 per cent, for inmates in State and Federal prisons. We use the ratio among jail inmates because it seems likely that the lower ratio in State and Federal prisons is driven by a larger proportion of offenders for violent and serious drug crimes relative to jails. The ratio of female to male offenders is likely to be even lower for these categories of crime.

[33] The inability of the model to fit the spike in the data at age 19 is not surprising. The profiles in the model are driven by the age profiles of wages, which are monotonic in the early working life.

[34] While the NLSY data are a cross-section of individuals between age 18 and 23 in 1980, the model simulation produces a panel of individuals. We construct a cross section for the simulated data, by drawing a random sample of 1000 individuals at each age between 18 and 23 taking care to exclude individuals sampled at an earlier age from the sampling population at later ages.

for residual heterogeneity.[35] For the simulated data, the additional control is the crime fixed effect χ. For the NLSY data we use variables for family background, race, ethnicity and geographic location (SMSA status). Column (5) also includes a dummy for whether the individual lives in the city centre. Again, the estimates are not statistically different across the two sets of data, and economically significant. The fall in the coefficient on the high school graduation dummy from column (1) to (2) is of interest in itself and reflects the fact that individuals with a higher residual (un-modelled) propensity to engage in crime (a higher χ) are less likely to graduate from high school. When the econometrician is unable to observe and fully control for this heterogeneity the estimated impact of high school completion on crime participation is upward biased.

Finally, column (2) in Table 5 reports a number of statistics for the benchmark equilibrium. Output and aggregate consumption (the net flow of consumable resources) are normalized to 100. The welfare criterion we employ is the permanent consumption level that would give an individual the same ex-ante (before ability, wealth and the crime fixed-effect are realized), expected (with respect to the idiosyncratic unemployment and victimization shocks) lifetime utility of a newborn into the stationary equilibrium.

The ratio between the victimization rate and the participation rate[36] in the first column in Table 5 is the average number of crimes committed by a criminal in one year. This ratio is 1.4 in the benchmark economy. It is important to keep in mind that one crime in the model yields an income equal to the average loss to the victim of one crime in the data. This is $728 current dollars in 1980.

4.2 Subsidizing High School Completion

The first experiment we carry out involves subsidizing high school completion. The dramatic increase in the U.S. prison population over the past twenty years has prompted a shift of interest, among both academics and policymakers, from tougher sentencing to other forms of intervention.[37]

[35] Column (4) is effectively the regression in column (4), Table 12 in Lochner and Moretti (2004) on the whole sample. We do not include the local unemployment rate among the regressors for lack of availability.

[36] The victimization rate is the number of crimes over the population out of jail. The participation rate is the number of criminals over the population out of jail.

[37] See, e.g., Donohue and Siegelman (2004).

A well-documented experiment, similar in spirit to a high school graduation subsidy, was indeed conducted by the Department of Labor and the Ford Foundation. The experiment - known as the Quantum Opportunities Program - was carried out on a small scale in two waves: a first one between 1989 and 1993 and a later one between 1995 and 2001. The program was targeted at adolescents from problematic schools. The experiment, appropriately randomized, offered learning support and cash incentives, from grade nine through to high school graduation, to students in the treatment group. It involved a 'salary' starting at $1 and rising to $1.33 per each hour of "activity" the student attended up to a ceiling, plus a bonus of $100 for each 100 program hours for completing activities.[38] An amount equal to the earned stipend was also deposited in an accrual account and paid to the enrollee conditionally upon completion of her high school degree. The total cost of the program was $3130 per student per year, of which $2150 represented the direct payment to the student and the remaining amount the cost of the resources (teaching support and equipment) the student had access to. Hahn, Leavitt, and Aaron (1994) report that the program had uneven effects in terms of reducing crime among participants relative to the control group. The effect was large and significant in a number of localities but negligible in others.

The experiments in this section consist in giving all targeted students attending high school a yearly subsidy equal to 8 per cent of average, labour earnings.[39]

4.2.1 Unconditional subsidy: general equilibrium

In this subsection, we investigate the effect of a subsidy paid to *all* individuals completing high school independently of their financial means.

Table 5 reports the changes in the education shares both in the aggregate and within each ability bin relative to the shares in the benchmark reported in Table 6. The subsidy increases the share of high school graduates in the population by one percentage point with only a marginal effect on the share of college graduates. The aggregate change is accompanied by increased sorting across ability bins. The share of high school graduates

[38] The maximum number of program hours was 750 divided between 250 hours of academic support, 250 hours of cultural and developmental activities and 250 hours of service activities, such as community service projects.

[39] This corresponds to the ratio between $ 3130 and average labour earnings in the data from 1995, one of the central years in which the Quantum Opportunities Program was run.

actually falls in the lowest ability bin while it increases in the others. The fall in the number of high school dropouts is particularly pronounced in bins 3 and 4.

The change in the education distribution can be understood as the outcome of two main forces. First, as it is apparent from Table 5, the individuals closer to the margin between attending high school or not in the benchmark are those in the highest ability bins who are held back by very low wealth. These individuals respond strongly to the subsidy. Second, the relative wage of high school dropouts increases as their relative supply falls. This change in relative prices has heterogeneous effects by ability on the return to high school completion. The fall in the education premium results in reduced incentives to graduate from high school for all agents, but proportionally more so for less able people.

Column (4) in Table 5 reports the same statistics as in the benchmark economy for our policy experiment. The victimization rate falls from 5.60 to 5.36 per cent, a 4.5 per cent reduction. Prison expenditure falls by a similar amount. Also the crime participation rate falls: from 3.95 to 3.89. The fall in the aggregate participation rate is the result of the increase in the proportion of high school graduates, who have a lower participation rate relative to dropouts, but also of the substantial fall in the participation rate among dropouts. The first effect is due to the general equilibrium increase in wages for high school dropouts which increases their opportunity cost of crime. The two effects more than offset the small increase in the average participation rate among high school graduates. This reflects the change in the composition of the pool of high school graduates. The average crime propensity of those individuals who switch as a consequence of the subsidy is marginally higher than the average crime propensity of the pool of high school graduates in the benchmark.

The improved education sorting results in an increase in both output and aggregate consumption of about 0.5 per cent. The effect on welfare, as measured by the ex ante consumption equivalent, is even more significant, nearing 2 per cent. By increasing the relative price of labour for high school dropouts and weakening the link between wealth at birth and selection into education, the subsidy provides insurance against the (ex-ante) uncertain ability and initial wealth draws. Since the ability shock is permanent and the initial wealth shock has a persistent effect through the education choice, the welfare benefits are large as they cumulate over the whole lifetime.

We experiment with a subsidy equal to twice and three times the original one. Each increase raises the aggregate share of high school graduates by roughly 0.5 percentage points, though at a mildly decreasing rate. The effect on crime, output and welfare is substantial. In particular, welfare increases by about 3 and 4 per cent relative to the benchmark in the case of a subsidy equal to 16 and 24 per cent of average labour income. The marginal benefit, in terms of crime reduction, increase in efficiency and welfare, is decreasing reflecting the progressive exhaustion of the benefits from improved sorting into education.

Note that while the total subsidy expenditure increases with the size of the subsidy up to a sizable 0.75 per cent of aggregate consumption in the benchmark, the labour tax rate is hardly affected as the increase in the tax base generates the necessary increase in revenue. The labour tax rate necessary to balance the budget is 27.3 per cent in the case of a subsidy equal to 24 per cent of average labour income. This is only marginally higher than its value of 27 per cent in the benchmark.

4.2.2 Unconditional subsidy: partial equilibrium

To understand the quantitative relevance of the general equilibrium effects discussed in the previous subsection, we now conduct the same experiment in partial equilibrium. Basically, we solve the problem of individual agents keeping all prices at their level in the benchmark economy. This type of experiment can be interpreted as a small scale intervention that has no effect on prices. We still let the aggregate crime rate adjust.[40] In fact, we are exactly interested in isolating the direct effect of the subsidy on the crime rate from its indirect effect due to the change in prices.

We restrict attention to a subsidy equal to 8 per cent of average labour income. Table 5 reports the change in the education shares relative to the benchmark. The share of high school graduates increases by 6 percentage points, six times its increase in general equilibrium. The increase affects all ability bins, though proportionally it is larger for higher ability bins.

Turning to column (2) in Table 5, the fall in the victimization rate is half as in general

[40] In this sense, it has to be interpreted as a local crime rate.

equilibrium despite an almost identical reduction in the aggregate participation rate. The fall in the average participation rate is mostly due to the change in the education distribution. The increase in the share of high school graduates almost fully offsets the smaller fall in the participation rate for dropouts and the larger increase in that of high school graduates, relative to the general equilibrium case. The substantially smaller fall in the victimization rate, though, reflects a higher average crime intensity in partial relative to general equilibrium. The increase in the relative price of high school dropouts in general equilibrium implies a much larger fall in the average number of crimes per criminal among high school dropouts[41] than in partial equilibrium. The magnitude of this general equilibrium effect is large enough to generate a nearly twice as large fall in the average crime intensity and victimization rate, relative to partial equilibrium. The welfare gain is marginally higher that in partial equilibrium.[42] The fact that real labour prices are unchanged benefits high school graduates and damages high school dropouts relative to the general equilibrium case. The number of high school graduates, though, is more than twice as much as that of dropouts.

4.3 Increasing the prison sentence

We now turn to the allocational and welfare effects of an increase in the prison term and compare them to the effects of the education subsidy analysed in the previous two section. A natural way to compare the two policy is to consider an increase in the prison sentence that achieves *the same reduction* in the victimization rate as the education subsidy.

Columns (7)-(9) are meant to be compared with columns (4)-(6). A prison term of respectively 19.6, 20 and 20.3 months achieves the same victimization rate as a subsidy equal to 8, 16 and 24 per cent of average labour income. The policy has basically no effect on the education distribution and on prices and output. Despite the same fall in the victimization rate, the total prison expenditure falls by less than in the case of the high school subsidy. The increase in the sentence length implies a higher stock of inmates for the same victimization rate.

[41]The average number of crimes among high school graduates in basically unchanged.

[42]We do not report aggregate output measures because we solve the problem for individual agents along their factor supply curves, hence off the production function.

Interestingly, the same fall in the victimization rate requires a larger fall in the number of criminals relative to the education subsidy. The intuition is apparent by inspecting the crime participation rates for drop outs and high school graduates. The participation rates fall similarly in both groups. Yet, contrary to the subsidy experiment there is neither an increase in the share of high school graduates, nor an increase in the relative price of high school dropouts. Both the general equilibrium and the composition effect, which contributed to a fall in the victimization rate in the case of the subsidy, are absent here. Welfare increases only marginally reflecting the reduced transitory income risk from property crime.

4.4 Small versus large interventions: the importance of equilibrium effects

A conditional education subsidy is described by two attributes: first, the transfer amount received by each agent; second, the number of agents who end up taking advantage of it. While the first attribute is fully determined by policy-makers, the second is not because not all eligible people are willing or able to take up the subsidy. Therefore the effects of alternative policy interventions depend not only on the transfer amount paid out, but on the number of participants as a share of total population. Even more interestingly, only a subset of individuals within the group of program participants is marginal to the policy, meaning that they change their education choices because of it.

One can define the 'direct' effect of a policy by the share of marginal individuals whose education (and crime) choices change because of the policy: this share becomes relatively less important as the scope of the policy grows and the eligible population becomes larger. At the extreme, large scale policies target mostly recipients who would go to high school anyway.

There is however also an 'indirect' effect of large scale policies, working through changes in the relative returns to education (and opportunity cost of crime). Interventions which affect many sections of the population can shift enough people into education so that relative marginal returns are affected: in turn, this has an effect on the ability composition in each education group, as more able people crowd-out less able ones and more sorting takes place. Therefore it seems natural to ask what is the scale of an intervention for which sizeable equilibrium effects are present.

Tables 12, 13 and 14 offer some insights on the aggregate effects of policies with different transfers and different target populations. In each table we report the results of the general equilibrium experiment for a given transfer size, but different degrees of means-testing (namely, the 10th, 30th, 50th, 70th and 90th percentiles of the initial wealth distribution). The three tables present results for increasingly large transfer amounts (from 8 per cent of average labour earnings up to 3 times as much). All experiments show a strong positive welfare effect of education policies.

Regarding the relationship between means-testing and policy effectiveness, we notice that half of the gains in aggregate consumption, welfare and crime reduction are already present in experiments targeting the bottom 10% of the wealth distribution. As might be expected, policies with looser means-testing criteria offer only small additional gains because much of the transfers are made to infra-marginal individuals. Moreover, looking at the price of human capital of different types, we notice that equilibrium effects are already present for policies means-tested to low wealth. Figures 6, 9 and 12 report the equilibrium price changes for different subsidies at different levels of means-testing. They confirm that policies targeted to the lowest section of the wealth distribution are quite capable of triggering sizeable equilibrium effects. Figures 5, 8 and 11 show the effects of different policies on aggregate consumption, welfare and prison costs and confirm the strong gains associated to education policies, mostly reaped at low levels of means-testing. Finally figures 4, 7 and 10 shed some light on the cost of education policies as a share of aggregate consumption in the benchmark: remarkably, means-testing at 10% or 30% implies total costs of transfers and prison expenditures comparable to the ones in the benchmark, despite a large efficiency and welfare gain. In summary, all these experiments indicate a trade-off between transfer size and scope of interventions: comparing tables 12 and 14 it seems clear that policies with stricter means-testing, but larger transfers, are preferable to policies with looser means-testing but smaller transfers. The per-dollar effectiveness of the policy is much stronger at the lower end of the wealth distribution and, importantly, equilibrium effects seem to appear early.

4.5 Introducing a 'control group': A 50/50 randomisation of the Subsidy

Leaving aside any equilibrium effects, the most interesting piece of information regarding a subsidy experiment is how effective it is in reducing crime among those who receive it. In order to make a statement regarding such change one should be able to compare the crime rates of two groups which are 'ex-ante' identical in every respect but the subsidy.

To this purpose we perform a means-tested subsidy experiment with the simple variant that, among all eligible individuals, only a randomly chosen 50% actually receive the offer of a subsidy (with the remaining 50% not being offered anything or, in the language of the program evaluation literature, not being 'treated'). We then use the 'untreated' 50% of the eligible population as a 'control' group to assess the effectiveness of the policy. In particular we evaluate its impact on average criminal behavior over the full life-cycle of 'treated' individuals.

In the experiment we consider a conditional high-school subsidy equal to 8% of average labor earnings (the monetary equivalent of the 'Quantum Opportunity Programme' transfer). We focus on a means-tested subsidy for which eligibility is restricted to those in the lowest third of the wealth distribution at birth (interpreted as family resources accessible to youths). The experiment is run in partial equilibrium and the group of eligible individuals is followed over their entire life cycle.

We present results comparing average crime intensity (total number of crimes) over the life cycle for all eligible people, both those randomized in or out, and document facts about selection in terms of underlying heterogeneity. However, given the voluntary take-up of the policy, the treatment group includes two types of people: those who choose to study and take advantage of the subsidy and those who don't.

Table 15 documents policy take up in the treatment group: roughly two thirds of eligible individuals take advantage of the subsidy. When we compare the average ability percentile of those taking up the subsidy, in table 17, it is evident that mostly people with high ability take advantage of the transfer, with the average ability rank of the 'treated' being almost 4 times as large as that of the eligible individuals who *choose* not to go to high school.[43]

We use simulated data from the randomisation experiment to assess the policy impact on life-cycle crime behaviour by program status. Table 19 reports the average number of crimes committed in each period by individuals in different program groups. The most common comparison in the evaluation literature is between people who take up the program and those who are randomised out: carrying out this simple comparison in our experimental sample suggests a striking difference in long-term criminal behavior between these two groups, with program participants ('treated' who take up the subsidy) committing roughly half the number of crimes as the control group (0.055 crimes per capita versus 0.093). This corresponds to a drop in life-cycle crime of about 41%, by any standards a very large effect for a policy.

Such strong result, however, is partly due to the fact that people who take up the subsidy have higher ability as discussed above. In fact crime intensity for 'treated' people who chose not to take up the subsidy is substantially higher. Computing a weighted average of the *overall* effect using all treated people (those who take up and those who don't) gives an average crime intensity of 0.091, which suggests a drop in crime intensity of only 2.3% in each period of life. This simple point illustrates that selection on unobservables, like ability, can lead to overestimating the impact of a given policy intervention because the effects of a policy tend to concentrate in those subgroups of the population which respond strongly to it.

However, the 2.3% drop in crime intensity is only a lower bound of the 'true' effect because it compares people who take up the program to the entire 'control' group, rather than to the subset of 'ex-ante' similar people within that 'control' group. Carrying out the correct comparison suggests a 'true' crime reduction of roughly 5.3% over the life-cycle, a non-negligible impact.[44] These results suggest that a targeted small subsidy policy can reduce crime intensity in the relevant population by more than 5% in each period of life, with the benefit being spread over a long time horizon corresponding to the life cycle of

[43] Interestingly, no selection with respect to heterogeneity in crime fixed effects is present.

[44] Assuming that the crime intensity of the treated individuals who reject the policy is similar to that of similar individuals who were randomised out, and using the take up shares within the treatment group, we can easily compute the crime intensity measure for the subset of the control group which is comparable to the program participants (that is, select only those individuals with characteristics similar to those who took up the program). The resulting crime intensity for this group is 0.058.

the 'treated'. The effect is significant considered the relatively small size of the transfer. The results also indicate that means-tested policies can still have an uneven effect over the target population, with substantial selection taking place along unobservable dimensions.

5 Conclusions

We develop and calibrate a structural, life-cycle model with heterogeneous agents and optimal education, crime and saving decisions. We use the model to study property crime and compare two alternative sets of policies: subsidies for high school completion and increases in prison sentences. We find that, given the same target in crime reduction, an unconditional subsidy to high school completion has large efficiency and welfare gains which are absent in the case of increases in prison sentences. A large fraction of these gains can be achieved by aiming the subsidy at individuals in the lowest end of the wealth distribution. The framework can easily be extended to allow for other interventions such as wage subsidies, income tax credits and other redistributive policies.

When looking at the relative importance of equilibrium effects in determining the final outcome of an intervention our results indicate that changes in marginal returns are already apparent for relatively low levels of means-testing and tend to reinforce the education policy effects. This is in stark contrast to pure punishment policies for which no equilibrium effects are present. In the case of education policies, sizeable equilibrium effects are present despite a very conservative parameterization of the aggregate technology function: this is due to the presence of heterogeneity in the population which induces sorting in equilibrium. This result is interesting because it suggests that, even under technology parameterizations which allow for relatively high substitutability, heterogeneity can interact with policy interventions resulting in significant equilibrium effects.

We also find that, given a limited budget, the most effective policies are those targeting individuals from the poorest backgrounds with higher transfer sizes, rather than small transfer policies targeting large sections of the population.

Finally, through a randomisation of the subsidy, we provide partial-equilibrium evidence that education policies have very uneven effects on different groups of the target population, even when aimed only at the wealth-poor. In particular we find that, because of endogenous selection into policy participation, standard comparisons of 'treatment' and 'control' groups can be misleading and lead to an overestimate of the true effect of a policy. However, when appropriately controlling for selection, our experimental results still show significant effects of an education subsidy over life-cycle crime intensity for the 'treated' group.

References

Abraham, A. (2001): "Wage Inequality and Education Policy with Skill-biased Technological Change in OG Setting," Ph.D. thesis, Universitat Pompeu Fabra.

Attanasio, O. (1999): "Consumption," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1B. North Holland, Amsterdam.

Barillas, F., and J. Fernandez-Villaverde (2007): "A Generalization of the Endogenous Grid Method," *Journal of Economic Dynamics and Control*, 31, 2698–2712.

Becker, G. (1968): "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76, 169–217.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116(2), 705–746.

Carroll, C. D. (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 91, 312–320.

Cooley, T. F. (1995): *Frontiers of Business Cycle Research*. Princeton, N. J.: Princeton University Press.

Cozzi, M. (2006): "Hard Drug Addiction, Drug Violations and Property Crime in the US," Mimeo, Queen's University.

Cunha, F., J. Heckman, and S. Navarro (2004): "Counterfactual Analysis of Inequality and Social Mobility," Mimeo, University of Chicago.

Daly, M., O. Jackson, and R. Valletta (2007): "Educational Attainment, Unemployment and Wage Inflation," *Federal Reserve Bank of San Francisco Economic Review*, pp. 49–61.

Domeij, D., and J. Heathcote (2003): "On The Distributional Effects Of Reducing Capital Taxes," Mimeo.

Donohue, J., and P. Siegelman (2004): "Allocating Resources among Prisons and Social Programs in the Battle against Crime," *Journal of Human Resources*, 39, 958–979.

Fella, G. (2009): "A Generalized Endogenous Grid Method for Non-Concave Problems," Mimeo.

Flinn, C. (1986): "Dynamic Models of Criminal Careers," in *Criminal Careers and 'Career Criminals'*, ed. by A. Blumstein, J. Cohen, J. A. Roth, and C. A. Visher, vol. II. National Academy Press, Washington D.C.

Freeman, R. (1999): "The Economics of Crime," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 3C, chap. 52. Elsevier Science Publishers.

Gallipoli, G., C. Meghir, and G. Violante (2009): "Equilibrium Effects of Education Policies: A Quantitative Evaluation," Mimeo, University College London.

Gould, E., B. Weinberg, and D. Mustard (2002): "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997," *Review of Economics and Statistics*, 84(1), 45–61.

Grogger, J. T. (1998): "Market Wages and Youth Crimen," *Journal of Labor Economics*, 16, 756–791.

Hahn, A., T. Leavitt, and P. Aaron (1994): *Evaluation of the Quantum Opportunities Program: Did the Program Work?* Brandeis University, Heller Graduate School, Waltham, MA.

Heathcote, J., K. Storesletten, and G. Violante (2008): "The Macroeconomic Implications of Rising Wage Inequality in the US," Mimeo.

Heckman, J., L. Lochner, and C. Taber (1998): ""Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," *Review of Economic Dynamics*, 1, 1–58.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch, vol. 1, chap. 7, pp. 307–458. Elsevier.

Imrohoroglu, A., A. Merlo, and P. Rupert (2000): "On the Political Economy of Welfare Redistribution and Crime," *International Economic Review*, 41, 1–25.

——— (2004): "What Accounts for the Decline in Crime," *International Economic Review*, 45, 707–729.

Kaplan, G., and G. L. Violante (2009): "How Much Insurance in Bewley Models?," Mimeo, NYU.

Lee, D. (2001): "An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice," Ph.D. thesis, University of Pennsylvania.

Lee, D., and K. I. Wolpin (2004): "Intersectoral Labor Mobility and the Growth of the Service Sector," PIER working paper 04-36.

Levitt, S. (1996): "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," *Quarterly Journal of Economics*, 111, 319–352.

——— (1997): "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, 87, 270–290.

Lewis, G. L. (1989): "Trends in Student Aid: 1963-64 to 1988-89," *Research in Higher Education*, 30, 547–561.

Lochner, L. (2004): "Education, work and crime: a human capital approach," *International Economic Review*, 45(3), 811–843.

Lochner, L. J., and E. Moretti (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *American Economic Review*, 94(1).

Machin, S., and C. Meghir (2004): "Crime and Economic Incentives," *Journal of Human Resources*, 39, 958–979.

Merlo, A., and K. I. Wolpin (2009): "The Transition from School to Jail: Youth Crime and High School Completion Among Black Males," PIER Working Paper No. 09-002.

NCHS (1997): *US Decennial Life Tables for 1989-1991.* Hyattsville:MD.

Pastore, A. L., and K. Maguire (eds.) (various years): *Sourcebook of criminal justice statistics.* U.S. Department of Justice, Bureau of Justice Statistics, Washington, DC.

Raphael, S., and J. Ludwig (2003): "Prison Sentence Enhancements: The Case Project Exile," in *Evaluating Gun Policy: Effects on Crime and Violence*, ed. by J. Ludwing, and P. Cook, chap. Chapter 7. Brookings Institutions, Washington, DC.

Taggart, R. (1995): *Quantum Opportunities Program.* Opportunities Industrialization Centers of America, Philadelphia.

Tobacman, J. (2009): "The Effect of the 2001 Tax Rebates on Crime," Mimeo.

Valletta, R. (2005): "Rising Unemployment Duration in the United States: Causes and Consequences," Mimeo.

Wolff, E. N. (2000): "Recent trends in wealth ownership, 1983-1998," working paper.

Year	High School Dropouts	High School	College
1977	0.23	0.41	0.36
1978	0.23	0.41	0.36
1979	0.23	0.41	0.37
1980	0.22	0.41	0.37
1981	0.22	0.41	0.37
1982	0.22	0.41	0.38
1983	0.21	0.40	0.38

Table 1: Shares of different types of human capital by year. CES human capital aggregation.

Age group	Unemployment duration (months)	Unemployment rate (per cent)
[16 – 19]	2.4	19
[20 – 24]	2.9	12
[25 – 34]	3.2	7
[35 – 54]	3.4	4.8
[55 – 64]	3.5	3.9

Table 2: Age-dependent unemployment shocks. Source: Valletta (2005) and Daly, Jackson and Valletta (2007).

Education group	Gradient (S.E.)
HSD	.36 (.06)
HSG	.54 (.03)
CG	.89 (.09)
pooled	.71 (.02)

Table 3: Estimated ability gradient. Sample 2: Wage = CPS-type

Parameter	Value	Moment to Match
$\bar{j}$	79	Maximum lifespan after labor market entry
j_r	50	Maximum years of working life
λ_j	-	Survival rates
β	0.958	Wealth-Income ratio excluding top 5%
CRRA coefficient	1.5	Attanasio (1999)
$\overline{a}$	-.75	Fraction of households with net worth ≤ 0
f_{HS}	.01	Direct tuition cost of High School
f_{CG}	.97	Direct tuition cost of College
α	.35	Capital share in total output
ρ	.68	From Gallipoli, Meghir, and Violante (2009)
σ_0	.22	From Gallipoli, Meghir, and Violante (2009)
σ_1	.41	From Gallipoli, Meghir, and Violante (2009)
δ	.065	Depreciation rate
p	.164	Pension replacement rate
t_l	.27	Labor income tax
t_K	.40	Capital income tax
α	.058	Average loss from property crime
π_p	.057	Probability of conviction
Sentence length	19 months	Average completed sentence
$\overline{c}$	.38	Victimization rate
χ	$U(-3.55, .17)$	Share of recividivists and HS dropouts among criminals

Table 4: Value of Parameters Calibrated in Benchmark

Dependent variable: real log hourly earnings ($1992)			
	Less Than HS Coefficient (Std. Err.)	High School Coefficient (Std. Err.)	College Coefficient (Std. Err.)
age	0.26	0.41	0.67
	(0.133)	(0.058)	(0.101)
age^2	-0.01	-.013	-.021
	(0.005)	(0.002)	(0.004)
age^3	1.5e-4	1.e-4	3.e-4
	(8.3e-5)	(4.e-5)	(6.e-5)
age^4	-9.7e-7	-1.e-6	-1.6e-6
	(4.9e-7)	(2.e-7)	(3.7e-7)
Intercept	-0.507	-2.23	5.12
	(1.232)	(0.533)	0.967

Table 5: Age polynomials' coefficients

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Aggregate rates
HSD	86	25	11	3	1	25
HSG	13	70	78	73	54	57
CG	1	5	11	24	45	17

Table 6: Shares of workers in different education groups by ability (AFQT89). Values are grossed-up in equal proportions to replicate the aggregate education shares observed for workers between 1977 and 1983 in CPS March supplement.

	Age 16		Age 18	
	Students	Workers	Students	Workers
Ability bin 1	1.96	1.05	1.11	1.68
Ability bin	1.49	0.16	1.05	1.16
Ability bin 3	1.29	0.01	1.08	0.92
Ability bin 4	1.21	0	1.10	0.80
Ability bin 5	1.15	0	1.17	0.50

Table 7: Average wealth (as a share of benchmark average) of students and workers at the ages of the education choices.

	Model data		NLSY data		
	(1)	(2)	(3)	(4)	(5)
HS Graduate	-5.85	-4.65	-6.43	-6.48	-5.63
	(0.9)	(0.84)	(2.20)	(2.23)	(2.25)
AFQT pct. (1-99)	0.07	0.05	0.04	0.01	0.06
	(0.01)	(0.01)	(0.03)	(0.04)	(0.04)
Crime fixed eff.		8.00			
		(0.27)			
Controls:					
Age	x	x	x	x	x
Family background				x	x
SMSA status				x	x
Ethnicity/race				x	x
Central city					x

Notes: Each row represents a separate OLS regression. The dependent variable is a dummy variables equal to one if the person participated in property crime. Unit of observation is individuals age 18-23 in the model data and males ages 18-23 in 1980 in the NLSY data. All coefficient estimates are multiplied by 100. In column (4), "Family background" includes: parents' highest school grade completed, whether or not the person lived with both natural parents at age 14, whether the person was born to a teenage mother.

Table 8: Linear probability regression on model-generated and NLSY data

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Aggregate
HSD	+6	+1	-10	-3	-1	-1
HSG	-6	-1	+10	+2	0	+1
CG	0	0	0	+1	+1	0

Table 9: Differences with respect to benchmark (absolute changes) in shares of workers in different education groups by ability (IQ test) bin, given a non means-tested high school subsidy (general equilibrium).

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Aggregate
HSD	-13	-7	-11	-3	-11	-7
HSG	+13	+7	+10	+2	54	+7
CG	0	0	+1	+1	+1	0

Table 10: Differences with respect to benchmark (absolute changes) in shares of workers in different education groups by ability (IQ test) bin, given a non-means tested high school subsidy (pe).

	Bench.	Subsidy PE	Subsidy GE			Prison		
HS subsidy	0	0.1	0.1	0.2	0.3	0	0	0
Prison sentence (months)	19	19	19	19	19	19.6	20	20.3
Crime victimization (%)	5.60	5.47	5.36	5.25	5.17	5.36	5.25	5.17
Crime participation (%)	3.95	3.90	3.89	3.88	3.87	3.86	3.82	3.79
Crime part. HSD (%)	6.40	6.29	5.98	6.02	6.08	6.32	6.22	6.16
Crime part. HSG (%)	3.92	4.17	4.06	4.07	4.06	3.80	3.77	3.75
Output	100.0	-	100.5	101.0	101.3	100.0	100.0	100.0
Agg. consumption	100.0	-	100.6	101.2	101.5	100.0	100.0	100.0
Welfare	100.0	102.2	101.9	103.3	104.3	100.1	100.1	100.1
Prison expenditure† (%)	0.27	0.26	0.25	0.25	0.24	0.26	0.26	0.26
Subsidy + prison exp.† (%)	0.27	0.51	0.48	0.72	0.96	0.26	0.26	0.26

† As a share of aggregate consumption in the benchmark.

Table 11: Subsidy and prison experiments. Subsidy as % of average labour income.

	Benchmark	MT10	MT30	MT50	MT70	MT90
HS subsidy (%)	0	8	8	8	8	8
Crime victimization (%)	5.60	5.50	5.44	5.44	5.40	5.40
Crime participation (%)	3.95	3.92	3.91	3.92	3.90	3.90
Output	100.00	100.20	100.32	100.42	100.44	100.52
Agg. consumption	100.00	100.26	100.40	100.52	100.54	100.63
Welfare	100.00	100.84	101.30	101.57	101.81	101.82
Prison expenditure† (%)	0.27	0.26	0.26	0.26	0.26	0.25
Transfers+Prison exp.† (%)	0.27	0.29	0.33	0.43	0.43	0.47
price LTHS	100.00	100.53	100.95	101.25	101.59	101.61
price HSG	100.00	100.01	100.00	99.95	100.02	99.88
price CG	100.00	100.02	99.97	99.86	99.89	99.79

† As a share of aggregate consumption in the benchmark.

Table 12: Effects of conditional HS transfer by threshold of means-testing (from 10th to 90th percentile of initial wealth). Subsidy: 8% of av. labour earnings.

	Benchmark	MT10	MT30	MT50	MT70	MT90
HS subsidy (%)	0	16	16	16	16	16
Crime victimization (%)	5.60	5.47	5.42	5.37	5.34	5.30
Crime participation (%)	3.95	3.92	3.93	3.92	3.90	3.89
Output	100.00	100.35	100.57	100.75	100.81	100.90
Agg. consumption	100.00	100.44	100.71	100.92	101.01	101.11
Welfare	100.00	101.38	102.22	102.65	102.90	103.13
Prison expenditure† (%)	0.27	0.26	0.26	0.26	0.25	0.25
Transfers+Prison exp.† (%)	0.27	0.31	0.43	0.52	0.61	0.68
price LTHS	100.00	100.91	101.76	102.26	102.72	103.10
price HSG	100.00	100.03	99.94	99.90	99.81	99.64
price CG	100.00	100.01	99.95	99.79	99.54	99.45

† As a share of aggregate consumption in the benchmark.

Table 13: Effects of conditional HS transfer by threshold of means-testing (from 10th to 90th percentile of initial wealth). Subsidy: 16% of av. labour earnings.

	Benchmark	MT10	MT30	MT50	MT70	MT90
HS subsidy (%)	0	24	24	24	24	24
Crime victimization (%)	5.60	5.48	5.36	5.31	5.28	5.24
Crime participation (%)	3.95	3.93	3.92	3.91	3.89	3.88
Output	100.00	100.46	100.65	100.98	101.10	101.13
Agg. consumption	100.00	100.58	100.81	101.20	101.33	101.38
Welfare	100.00	101.83	102.70	103.49	103.93	104.10
Prison expenditure† (%)	0.27	0.26	0.25	0.25	0.25	0.25
Transfers+Prison exp.† (%)	0.27	0.35	0.52	0.66	0.81	0.91
price LTHS	100.00	101.40	102.14	102.93	103.56	104.10
price HSG	100.00	100.01	99.78	99.87	99.70	99.43
price CG	100.00	99.96	99.78	99.63	99.27	99.13

† As a share of aggregate consumption in the benchmark.

Table 14: Effects of conditional HS transfer by threshold of means-testing (from 10th to 90th percentile of initial wealth). Subsidy: 24% of av. labour earnings.

program status	Number	% of eligible population
eligible, treated, take-up	1,241	35.5
eligible, treated, no take-up	520	14.8
eligible, not treated	1,740	49.7
Total	3,501	100.0

Table 15: Program Status: break-down of eligible population

ability	share taking-up subsidy
bin 1	10.5
bin 2	57.5
bin 3	89.6
bin 4	100.0
bin 5	100.0

Table 16: Share of take-up among treated by ability bin

program status	mean	S.D.
eligible, treated, take up	63.9	22.9
eligible, treated, no take up	18.6	14.0
eligible, not treated	50.2	28.0
Total	50.0	28.5

Table 17: Mean and S.D. of ability percentile by program status - 1 lowest,99 highest

program status	wealth at age 16
among treated, take up	0.25
among treated, no take up	0.13
among untreated	0.20
non eligible	1.68

Table 18: Initial wealth (at age 16, as a share of benchmark average) by program status

	Average number of crimes per period over the life cycle
among treated, take up	0.055
among treated, no take up	0.178
among untreated	0.093
non eligible	0.050

Table 19: Average number of crimes per person in each period of the life-cycle by program status

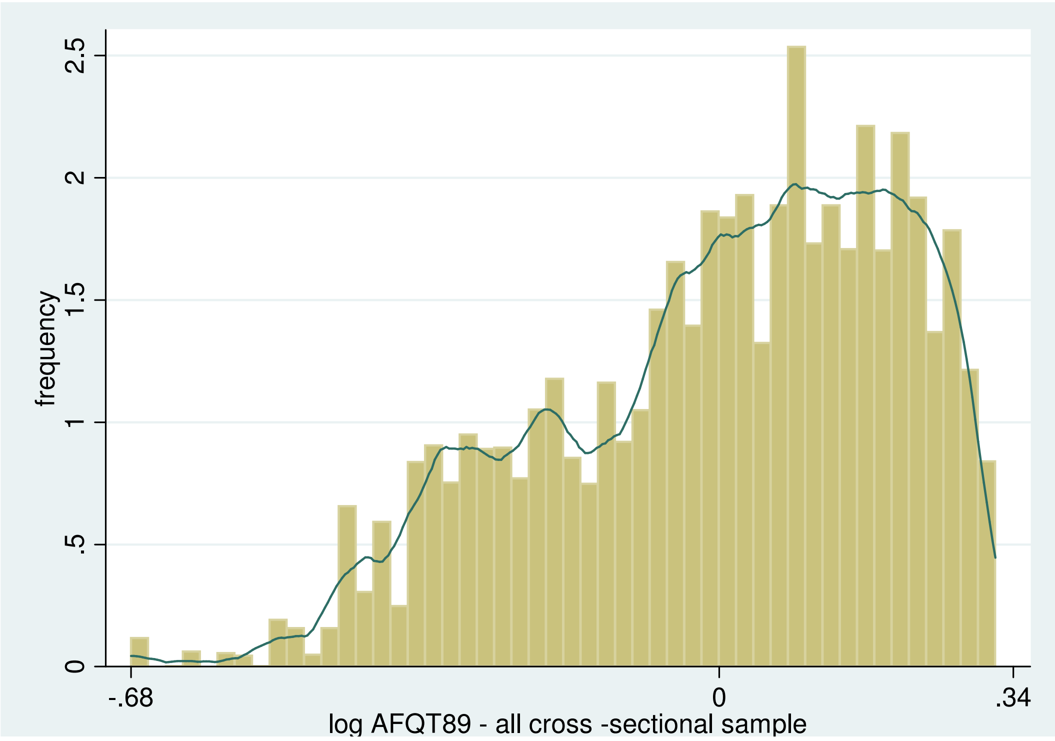

Figure 1: Distribution of ability (AFQT89 test) in a representative sample of the U.S. population.

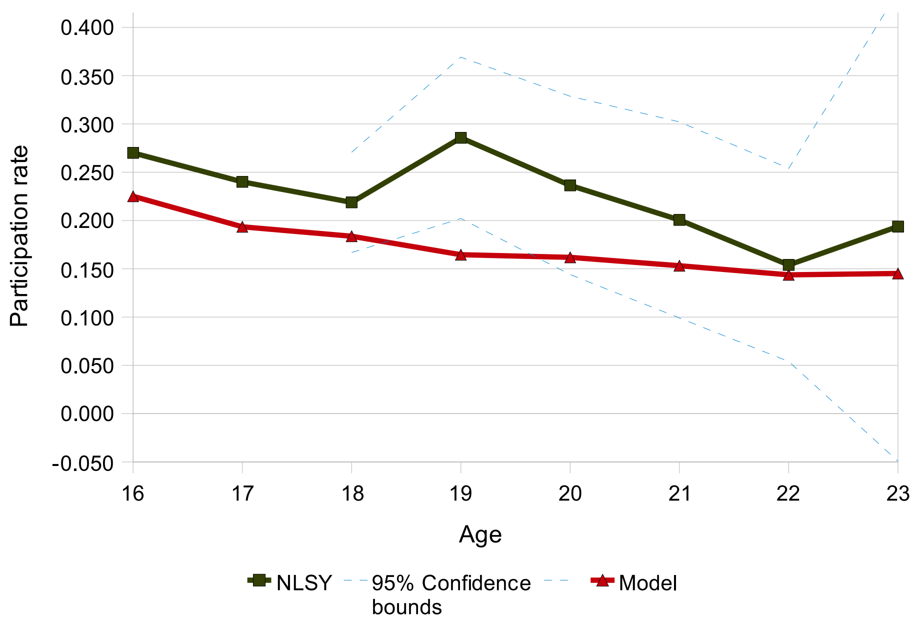

Figure 2: Crime participation by age: High School drop-outs

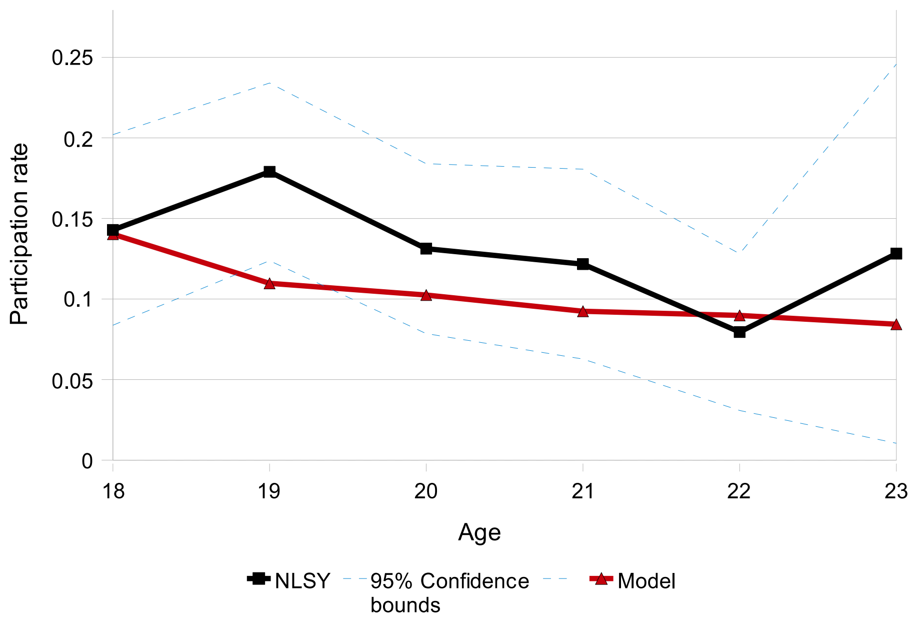

Figure 3: Crime participation by age: High School graduates

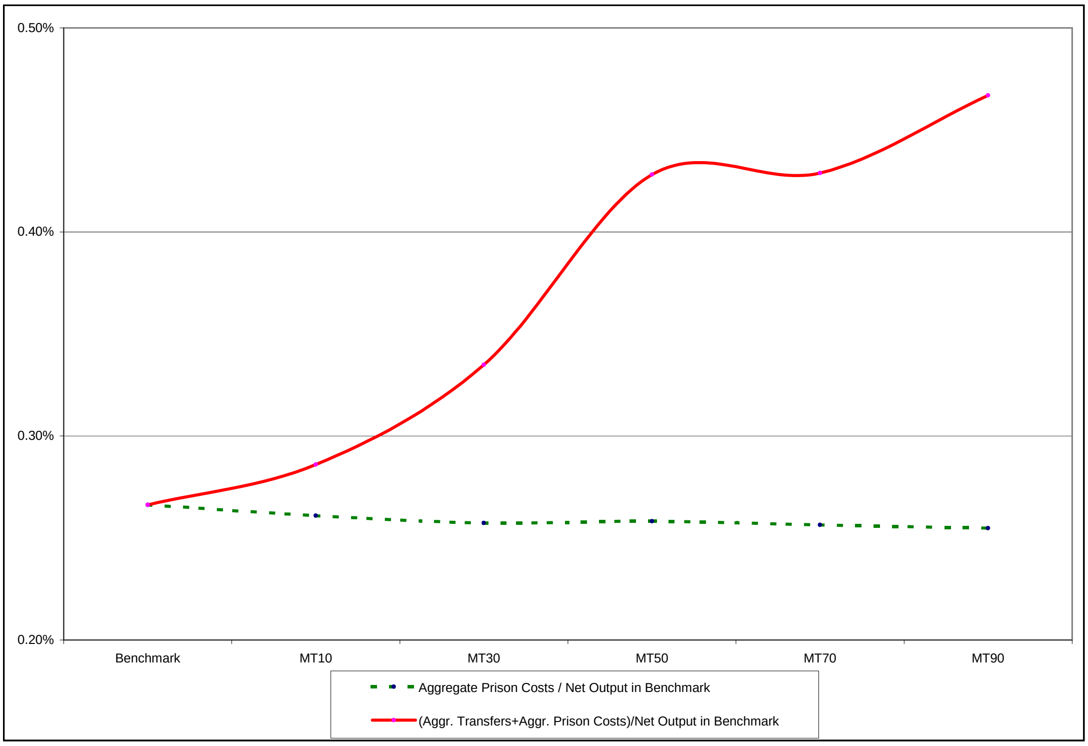

Figure 4: Costs of subsidy policy. Transfer size: 8% of av. labour earnings.

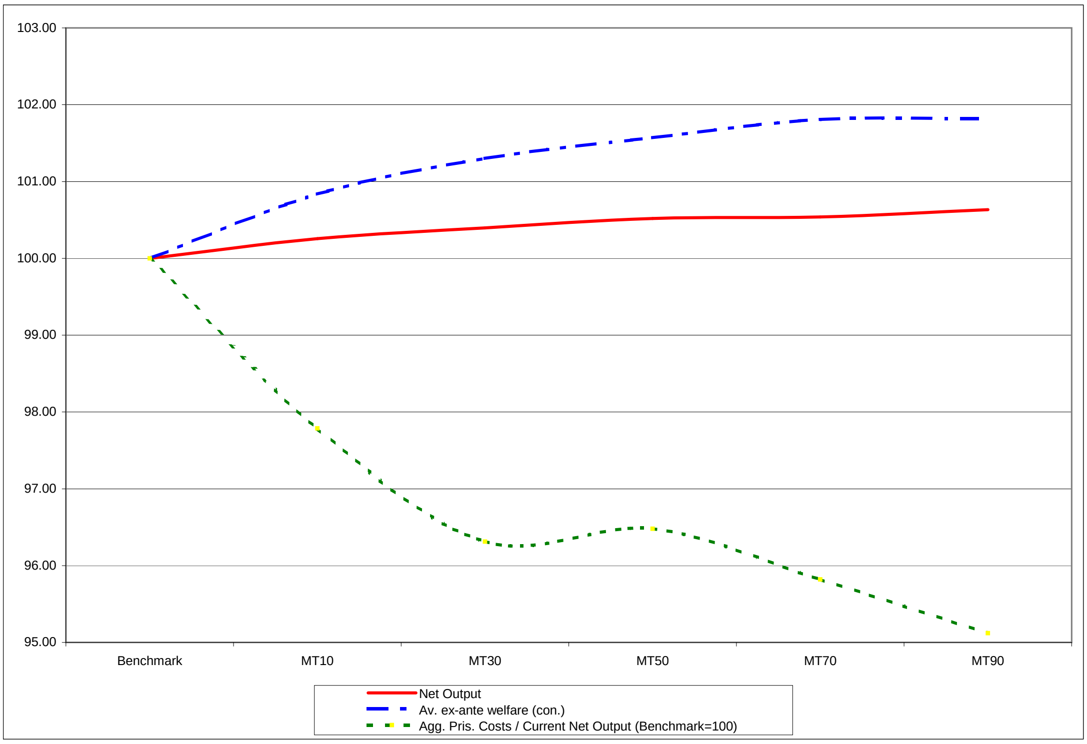

Figure 5: Aggregate effects of subsidy policy. Transfer size: 8% of av. labour earnings.

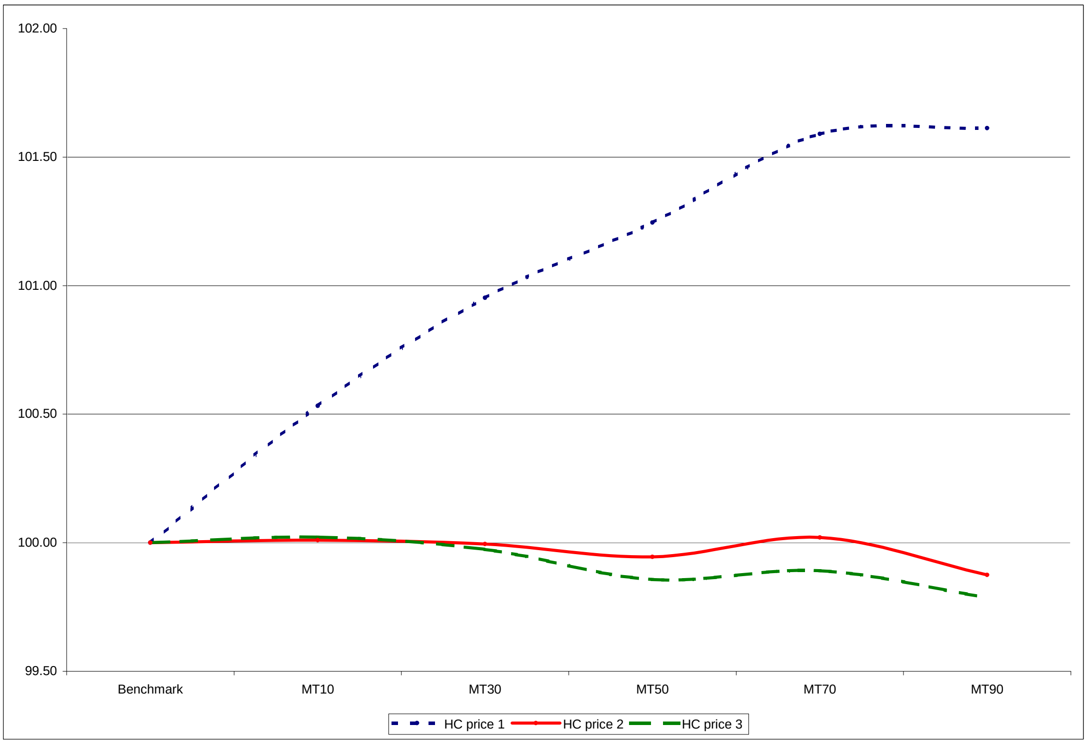

Figure 6: Equilibrium price changes induced by subsidy policy. Transfer size: 8% of av. labour earnings.

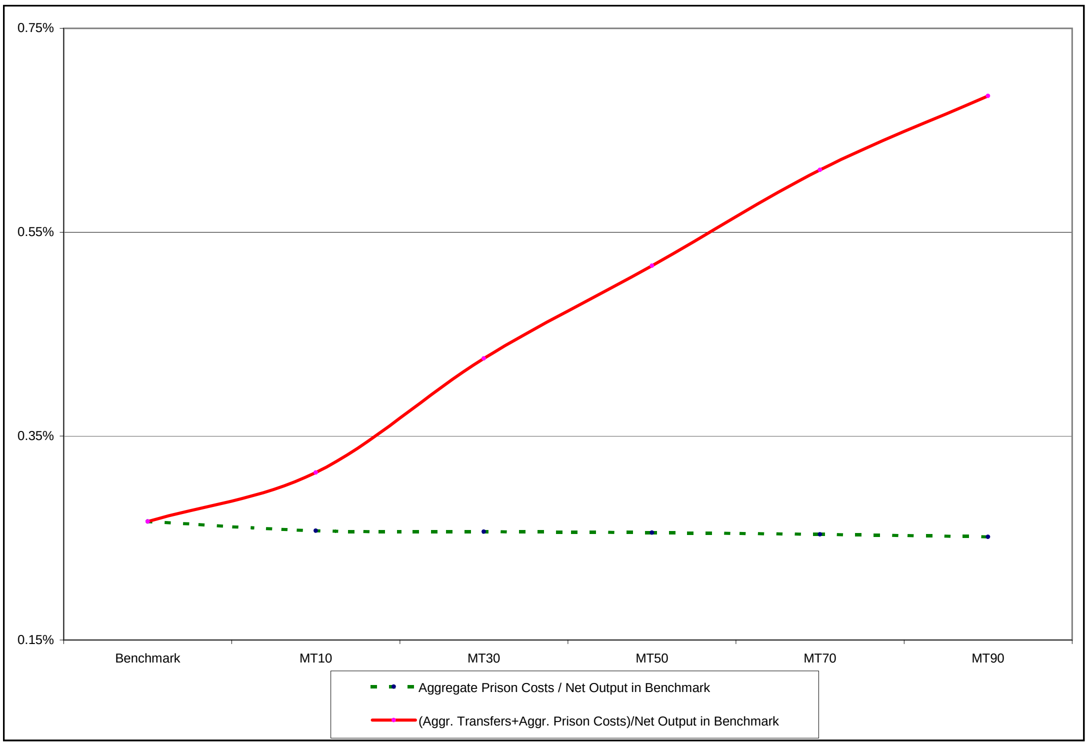

Figure 7: Costs of subsidy policy. Transfer size: 16% of av. labour earnings.

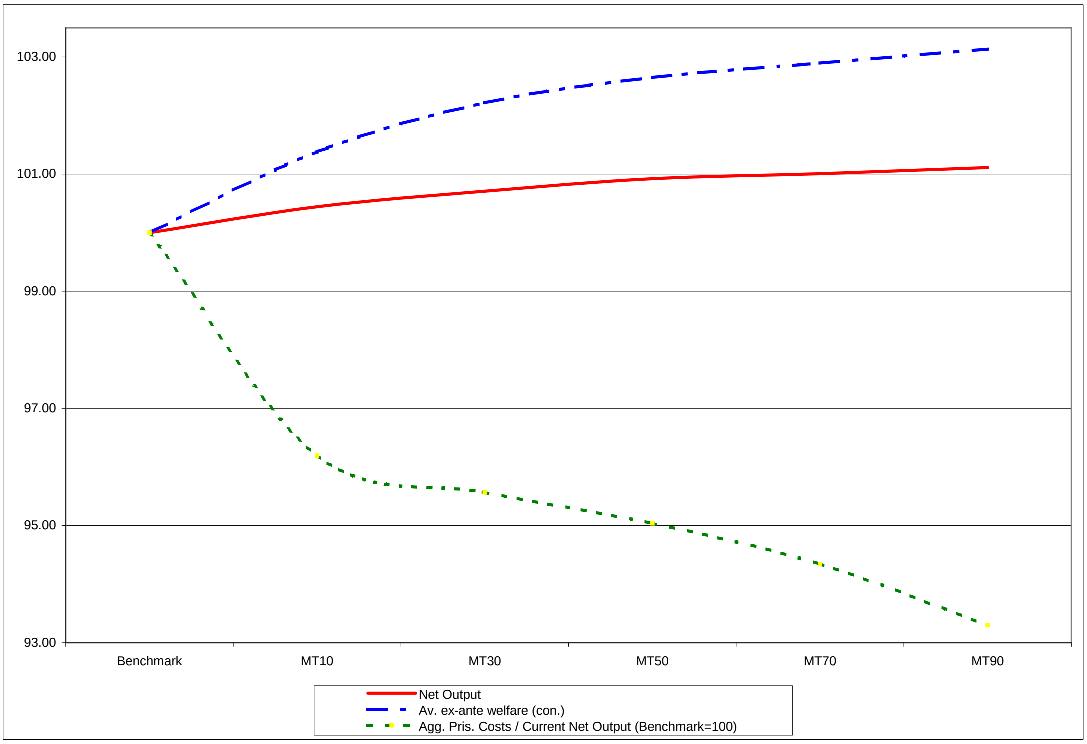

Figure 8: Aggregate effects of subsidy policy. Transfer size: 16% of av. labour earnings.

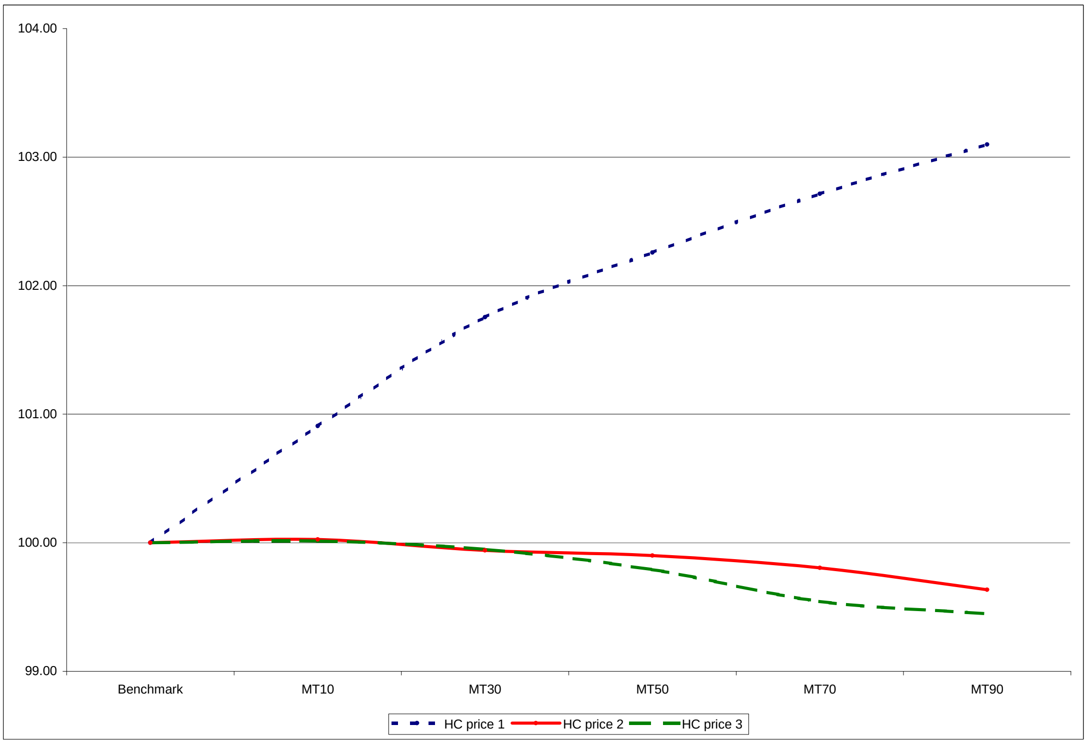

Figure 9: Equilibrium price changes induced by subsidy policy. Transfer size: 16% of av. labour earnings.

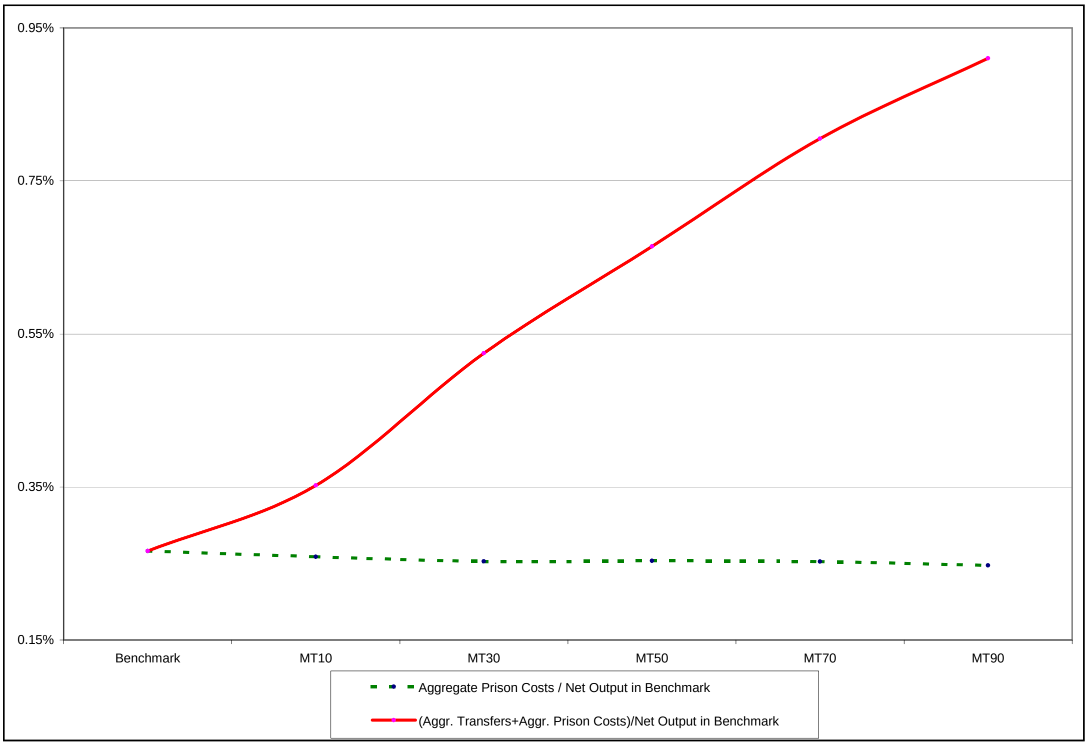

Figure 10: Costs of subsidy policy. Transfer size: 24% of av. labour earnings.

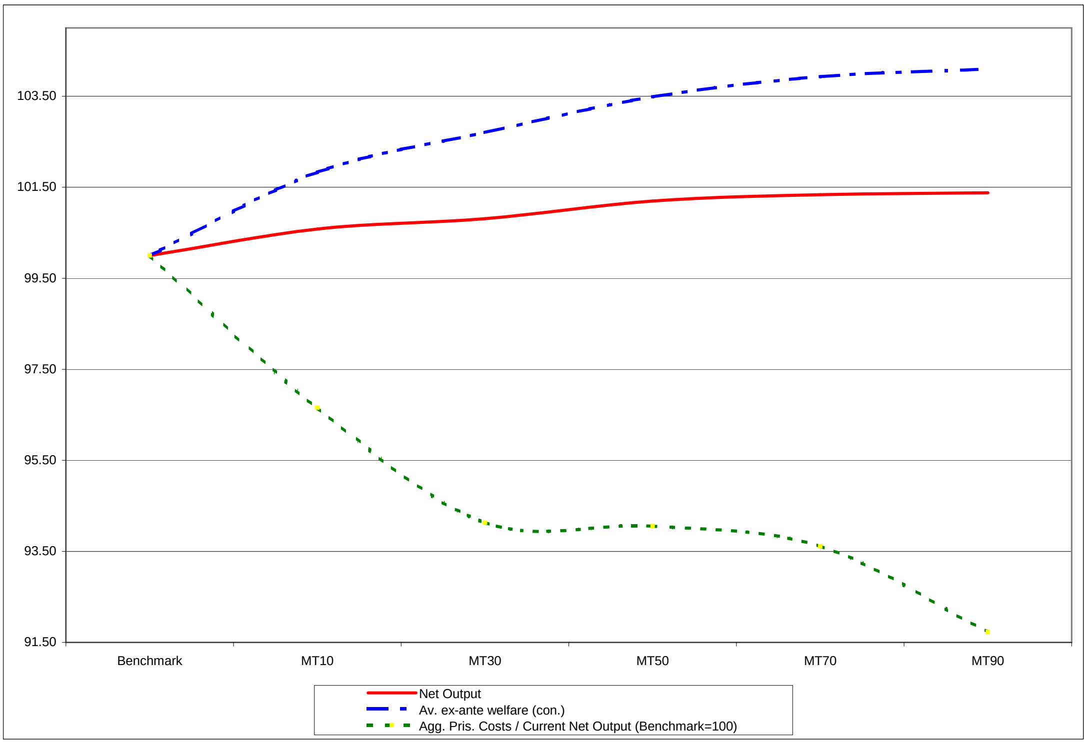

Figure 11: Aggregate effects of subsidy policy. Transfer size: 24% of av. labour earnings.

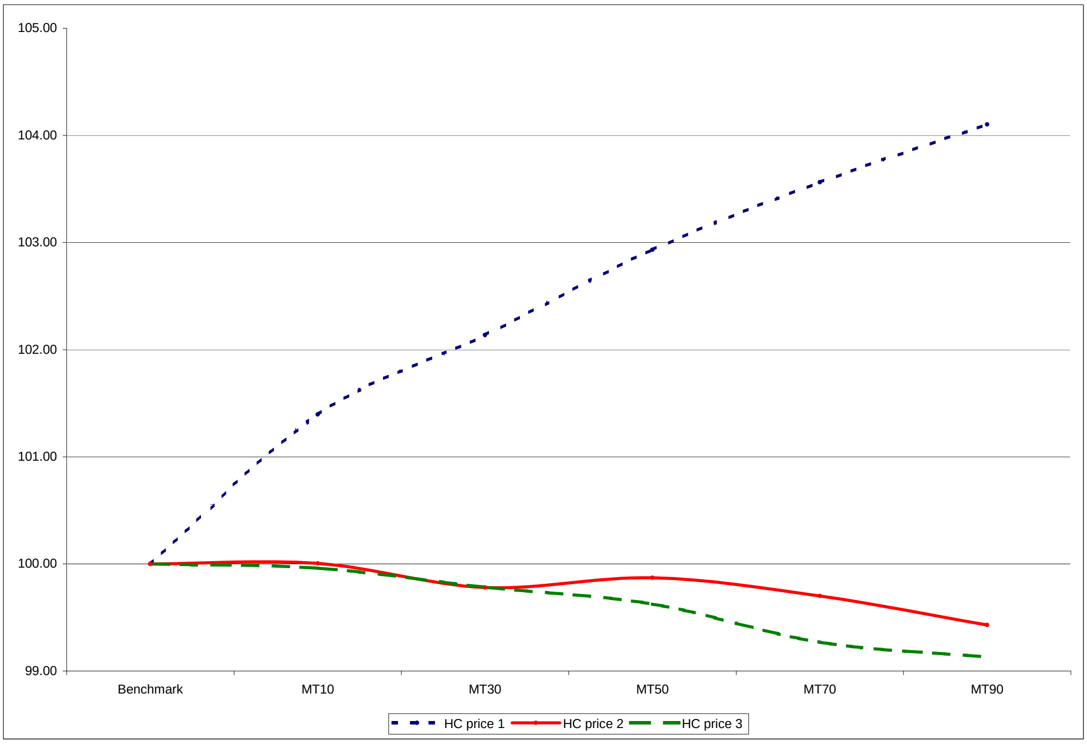

Figure 12: Equilibrium price changes induced by subsidy policy. Transfer size: 24% of av. labour earnings.