Appendix to “Education and Crime over the Life Cycle”

Giulio Fella and Giovanni Gallipoli

1 Appendix: 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 can take value s (student), n (worker out of jail) and p (worker in jail).

For a given set of government policies $\{pen, G, sub^e, t_l, t_k\}$, tuition fees $tuit^e$ and apprehension probability π_p, a *stationary recursive equilibrium* is a collection of (i) policy functions for consumption $c_j^x(z_j^x)$, saving $a_{j+1}^x(z_j^x)$, bequests b, education $\{i_H^s(z_0^s), i_C^s(z_{j_H}^s)\}$ and crime $\{\tau_j^n(z_j^n)\}$; (ii) value functions $\{V_j^x(z_j^x)\}$; decision rules $\{K, H_L, H_H, H_C\}$ for firms; (iv) prices $\left\{r, w^L, w^H, w^C\right\}$; (v) a victimization rate π_v; (vi) an average labor income $\overline{wh}$; (vii) time-invariant measures $\{\mu_j^s, \mu_j^n,\}$ and $\Gamma(a)$ that satisfy the following conditions.

1. Given prices$\left\{r, w^L, w^H, w^C\right\}$:
   - for $x = s, n$ the decision rules $\{c_j^x(z_j^x), a_{j+1}^x(z_j^x)\}$ and the value functions $V_j^x(z_j^x)$ solve respectively equations (13)-(15) for $x = s$, equation (16) for $x = n$ and $j < j_r, j \neq j_b$, equation (18) for $x = n$ and $j = j_b$, equation (19) for $j \geq j_r$;
   - the decision rule $a_{j+1}^p(z_j^x)$ satisfies equation (11) with $i^p = 1$ and the associate value function $V_j^p(z_j^p)$ solves equation (17);
   - the decision rule b solves equation (18);
   - the education decisions $\{i_L^s(z_0^s), i_H^s(z_{j_H}^s)\}$ solve equations (12) and (15);
   - the crime decision $\tau_j(z_j^n)$ solves equation (16).
2. Given prices$\left\{r, w^L, w^H, w^C\right\}$, input demands $\{K, H_L, H_H, H_C\}$ maximize profits for the representative firm

$$r = (1 - t_k)(F_K - \delta)$$

and

$$w^e = (1 - t_l) F_{H_e}, \text{ for } e \in \{L, H, C\}.$$

3. The asset market clears

$$K = \sum_{j,x} \int_{Z_j^x} a_{j+1}^x(z_j^x) d\mu_j^x.$$

4. The labor markets for each educational level clear[1]

$$H_e = \sum_{j<j_r} \int_{\{z_j^n : e=i\}} h_j(\theta, e) \left[1 - \pi_p \tau(z_j^n) - \pi_p f \tau(z_{j-1}^n)\right] d\mu_j^n, \text{ for } i \in \{0, 1, 2\}.$$

where the supply of labor on the right hand side of the above equation is made up only of individuals out of jail. In the calibrated model the prison term is strictly between one and two years. It follows that in the stationary equilibrium, the number of convicted felons in each age group j is the fraction $(1 - \pi_p \tau(z_j^n) - \pi_p f \tau(z_{j-1}^n))$ of workers that have not been arrested at age j or that, if arrested at age $j-1$, have not been released. To clarify notation, the argument z_{j-1}^n in the equation above is meant as a function of z_j^n such that z_{j-1}^n is identical to z_j^n with the only exception that asset holdings at age j satisfy $a_j = a_{j-1}^p(z_{j-1}^n)$.

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 $E = \sum_{j,i} \int_{\{z_j^s : e=i\}} sub^i \, d\mu_j^s$, for $i = \{L, H, C\}$, aggregate prison expenditure[2] $PRIS = \sum_{j<j_r} \int_{Z_j^n} m \, \pi_p \tau(z_j^n)) d\mu_j^n$ and aggregate pension expenditure $PENS = \sum_{j \geq j_r} \int_{Z_j^n} pen \, d\mu_j^n$.

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

[2] 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.

6. The victimization rate coincides with the crime rate

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

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

7. The average disposable labor income satisfies

$$\overline{wh} = \left(\sum_{j<j_r} \int_{Z_j^n} (1 - \pi_p \tau(z_j^n) - \pi_p f \tau(z_{j-1}^n)) d\mu_j^n \right)^{-1} \sum_e w^e H_e.$$

8. The distribution of wealth at birth $\Gamma(a_0)$ equals the distribution of bequests

$$\Gamma(a_0) = \int_{\{z_{j_b}^n : b(z_{j_b}^n) \le a_0\}} d\mu_{j_b}^n.$$

9. The vector of measures $\mu = \{\mu_0^s, ..., \mu_j^s; \mu_0^n, ..., \mu_j^n\}$ 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 $\{\chi, \theta, v, i^p, \varepsilon_j^e\}$ and the survival probabilities $\{\lambda_j\}$.

2 Appendix: Computation and calibration

Let $Z = \{\xi, \nu_1, \nu_2, \nu_3, \underline{a}, \underline{\chi}, \alpha, \beta, \rho_{\theta\chi}, \kappa, \bar{c}\}$ denote the set of calibrated parameters other than the utility cost of studying parameters $\{\psi^H(\theta), \psi^C(\theta)\}$. Given a guess for Z, $\{\psi^H(\theta), \psi^C(\theta)\}$ and the vector of equilibrium prices $\{r, w^L, w^H, w^C\}$ we calibrate the model in the following way.

1. We solve for the consumer decisions rules and value function and the representative firm factor demand functions.

2. We simulate the model up to the age of the college choice and solve for the values of $\{\psi^H(\theta), \psi^C(\theta)\}$ that match the enrolment rates in the data. Using the new values

as our new guess, we simulate again the economy up to the age of college choice and iterate on this procedure until the values of $\{\psi^H(\theta), \psi^C(\theta)\}$ converge.

3. We simulate the model at the remaining ages and compute the aggregate factor supplies. We compare the marginal products of the four factors to our guess for their prices. If the two differ by more than the specified tolerance, we adjust the guess for prices and solve again the problem, starting from point 1. until convergence (market clearing).

4. When factor prices have converged, we evaluate the loss function – the sum of squared deviations of the model from the data calibration moments – at the simulated model moments. We use a multi-dimensional optimization method to update the guess on Z and continue to iterate starting from point 1. above until convergence.

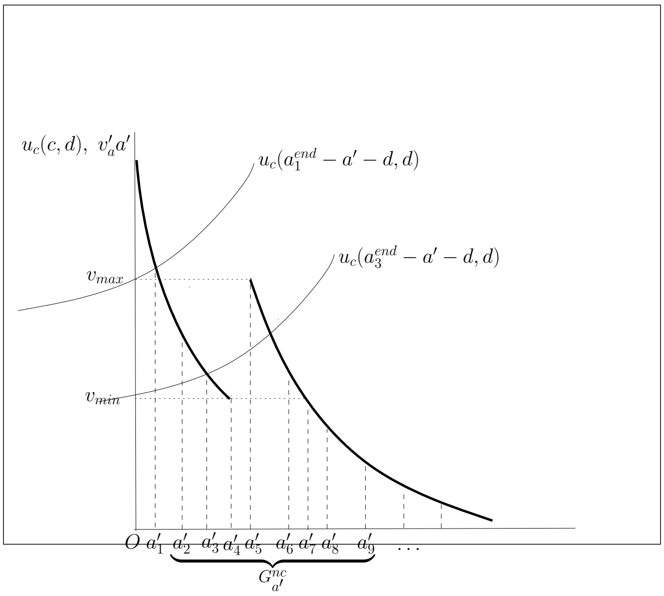
Concerning point 1. the decision rules and value functions point are computed using a generalized version of the endogenous grid method developed in Fella (2014). The method extends the original idea of Carroll (2006) to environments with non-convex choice sets.[3]

While the reader is referred to Fella (2014) for the details, we include here a brief sketch of the algorithm in the context of a simple problem.

Consider an agent with a two-period lifetime who derives intra-period utility $u(c, d)$ from consuming quantity c of a continuous good and quantity $d \in D = \{0, 1\}$ of a discrete good. The utility function satisfies the usual regularity conditions and, for simplicity, the Inada condition $u'(0, \cdot) = +\infty$. The relative price of the two goods is one. The agent has an initial endowment a of the continuous good. Both the (net) rate of return on storage and the agent subjective discount rate equal zero. There is no borrowing.

[3] Barillas and Fernández-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.

Figure 1: Solving for the conditional policy correspondence



The agent's problem in recursive form is

$$
\begin{aligned}
v(a) &= \max_{a' \in [0,a], d \in D} u(a - a' - d, d) + v'(a') \qquad (1)\\
v'(a') &= \max_{a'' \in [0,a], d' \in D} u(a' - a'' - d', d')\\
&\quad a \text{ given.}
\end{aligned}
$$

It follows that $v'(a') = u(a' - \hat{d}'(a'), \hat{d}'(a'))$ with $\hat{d}'(a') = \arg\max_{d' \in D} u(a' - d', d')$. The non-convexity of D, implies that, to the extent that $\hat{d}'(a')$ is not a constant, $v'(a')$ is neither concave nor differentiable and neither is the maximand and on the right hand side of (1).

Yet, Theorem 2 in Clausen and Strub (2012) implies that if, for given a, $(\hat{a}', \hat{d})$ is a maximum for (1) and $\hat{a}'$ is internal then $\hat{a}'$ satisfies the Euler equation

$$
u_c(a - \hat{a}' - \hat{d}, \hat{d}) = v'_a(\hat{a}'), \qquad \text{(EE)}
$$

as $v'_a(a')$ can jump up but not down.[4]

[4]This implies that the value of the Euler equation jumps up at discontinuities of $v'(a')$. Therefore a maximum cannot be located at a discontinuity.

Figure 1 plots the right and left hand sides of equation EE as a function of a' for a given value of d, The left hand side is plotted for two possible values of initial assets a. For given a, the intersection of the two curves is a candidate solution for the saving correspondence $a'(a|d)$ *conditional* on the given value of d. The contribution of Fella (2014) concerns how to solve for this "conditional" saving correspondence $a'(a|d)$.

In the standard approach, one fixes values for the endogenous state variable a at the beginning of the period and solves the Euler equation forward for the associated values of end-of-period wealth a'. Carroll's (2006) endogenous grid method (EGM), instead, fixes an ordered grid $G_{a'} = \{a'_1, a'_2, \ldots, a'_m\}$ for *end-of-period* assets a' and solve for the value of *initial wealth* a_i^{end} that satisfies EE for each $a'_i \in G_{a'}$. for each $a'_i \in G_{a'}$.[5] This approach is substantially faster as the Euler equation is often linear in consumption, hence in a, but non-linear (and in our case not even continuous) in a'.

Since, given the non-concavity of the problem, a local maximum is not necessary a global one, the algorithm modifies the standard EGM in the following way. First, it partitions the set of grid points for future assets $G_{a'}$ into a *non-concave region* $G_{a'}^{nc}$ in which the Euler equation is not sufficient for a global maximum for a' and its set complement. In terms of Figure 1, given the grid $G_{a'}$ and the derivative of the continuation value $v_a(a')$ it determines the *non-concave region* $G_{a'}^{nc}$ as the set of grid points for which $v'_a(a') \in (v_{min}, v_{max})$.[6] Secondly, for all a'_i in the non-concave region, the algorithm supplements EGM with a global maximization step.

More formally, given $G_{a'}$, $v'_a(a')$ for $a' \in G_{a'}$ and d

1. Determine the non-concave region $G_{a'}^{nc}$. Initialize the counters $i = 1$ and $l = 1$
2. Solve EE for a_i^{end} given a'_i using EGM.
3. If $a'_i \in G_{a'}^{nc}$ then
   - find the maximizer of the discretized maximand for $a = a_i^{end}$; i.e. solve for

$$a'_g = \arg\max_{a' \in G_{a'}^{nc}} u(a_i^{end} - a' - d, d) + v'(a').$$

[5] In terms of Figure 1, at the grid point a'_1, for example, the EGM solves for the value of initial assets a_1^{end} associated with the unique element of the family of upward sloping curves, indexed by initial wealth a, that intersects $v'_a(a')$ at point a'_1.

[6] In Figure 1, $G_{a'}^{nc} = \{a'_2, \ldots, a'_6\}$.

- if $a'_g \neq a'_i$, a'_i is not a global maximum. Move to the next grid point—$i = i+1$—and go to 2.

4. Store the solution pair (a_i^{end}, a'_i) as $(a_{i_l}^{end}, a'_{i_l}) = (a_i^{end}, a'_i)$. As long as a' is not the last grid point, set $i = i+1, l = l+1$ and go to 2.

5. Having solved for the conditional saving correspondence $\{a_{i_l}^{end}, a'_{i_l}\}$ on the endogenous collocation points $\{a_{i_l}^{end}\}$ solve for the conditional value function given d

$$v_{i_l} = u(a_{il}^{end} - a'_{il} - d, d) + v'(a')$$

6. Evaluate interpolating functions through (a_{il}^{end}, a'_{il}) and (a_{il}^{end}, v_{il}) at $a \in G_{a'}$ to obtain the conditional policy and value functions $a'(a|d)$ and $v(a|d)$ on the original grid $G_{a'}$.

7. Maximize $v(a|d)$ over d to obtain $d(a)$ and $v(a)$.

In a longer (possibly infinite) horizon case, having obtained $v(a)$ one would compute its partial derivative $v_a(a)$ and would work backwards.

Fella (2014) compares the accuracy and speed of the method to that of discretized value function iteration (VFI)—the most commonly chosen algorithm for non-concave, non-differentiable problems—using a saving problem with a discrete durable and a continuous non-durable choice. The discrete non-durable choice can take seven values, which implies a number of potential discontinuities larger than in the current model. He finds that the modified EGM algorithm has an accuracy, measured by the average Euler error (in base 10 log points) over a simulated history, in excess of -5 already with only 200 grid points for the continuous wealth variable. This is more than twice the accuracy of VFI[7] for the same number of grid points.[8]

[7]The average Euler error, rather than the supremum of the Euler errors, is the sensible accuracy measure in a model with discontinuities in the policy function, since, no matter how large the number of grid points, the probability of interpolating across a discontinuity goes to one as the length of a history increases. The Euler error when interpolating across the discontinuity is determined by the size of the jump in the function.

[8]In fact, the modified EGM with 200 grid points is still two orders of magnitudes more accurate, and 70 times faster, than VFI with 1000 grid points.

3 Appendix: College subsidy

The main policy focus of our analysis has been the effect of high school subsidization. Such focus naturally relates to the existing literature—reviewed in Lochner (2011)—on the effects of schooling on crime. In this literature a motivation for early intervention is that the majority of property crime is committed by people with relatively low education, and high school graduation has been proven effective in reducing crime.

In this appendix, instead, we analyze the effects of a policy that subsidizes college completion. In particular, we consider a transfer paid to all people who enroll and complete college. For comparability with the other policy experiments in the main body of the paper, we consider a college subsidy that achieves the same general equilibrium crime reduction as the high school subsidy equal to 8.8% of average earnings studied in the paper. The size of the college subsidy that achieves the targeted victimization rate of 5.2% is about 15.3% of average earnings.

Table 1 reports the effects (relative to the benchmark) of the college subsidy in both partial and general equilibrium. In general equilibrium the policy costs as much as the high school subsidy policy, but it induces a somewhat smaller welfare gain. Intuitively a college subsidy provides less insurance against ex ante uncertainty, relative to the high school subsidy, as it primarily affects high ability individuals marginal to the college choice.

To understand what drives the general equilibrium response, it is instructive to consider what the effect would be in partial equilibrium. The policy actually *increases* the crime rate by almost 0.3 percentage points in partial equilibrium as it induces a very large rise in college completion. The size of the subsidy, together with the relatively low cost of college attendance in 1980, implies that college education becomes not only free at the point of entry, but is associated to a non-trivial monetary transfer. Such large shift in college completion, at constant prices, increases income inequality and the average return from crime. In contrast, the high school subsidy actually *reduces* the crime rate by a similar amount already in partial equilibrium. Therefore, all of the crime reduction effects of the college subsidy are due to general equilibrium effects. These are even larger than in the case of the high school subsidy, as the college subsidy substantially increases

Table 1: College subsidy experiments. Subsidy as % of average labor income.

	Benchmark	COL Subsidy PE	COL Subsidy GE
	(1)	(2)	(3)
COL subsidy	-	15.3	15.3
Prison sentence (months)	19	19	19
Crime victimization (%)	5.6	5.9	5.2
Arrest rate L (‱)	5.9	7.1	5.5
Arrest rate H (‱)	2.7	3.4	2.7
L share of criminals (%)	48	51.5	47
Output	100.0	-	101.6
Agg. Consumption	100.0	-	102.1
Welfare	100.0	106.3	103.1
Prison expenditure†	0.30	0.31	0.28
Subsidy + prison exp.†	0.30	0.88	0.51
Price L	100.0	-	102.4
Price H	100.0	-	102.3
Price C	100.0	-	97.9

† As a share of aggregate consumption in the benchmark.

the human capital price not only for high school dropouts, but also for high school graduates. The resulting increase in earnings among the lowest two education groups raises the opportunity cost of engaging in crime for those agents who are most likely to commit crime. The difference between partial and general equilibrium is stark and highlights the importance of general equilibrium adjustments, which would be the only anti-crime justification for implementing such a policy.

Finally, one word of warning is necessary when assessing these results on the effects of a universal college subsidy. The direct costs of attending college have been steadily increasing over time, and continue to do so. As we mentioned above, a universal college subsidy of 15.3% (the one considered in this experiment) would have resulted, in 1980, in free college at the point of entry plus a yearly handout equal to roughly half the college cost. The same proportional subsidy in 2000, when college tuitions were more than double those in 1980, would instead have covered only between half and two thirds of the direct cost of college.

References

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

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

Clausen, A., and C. Strub (2012): "Envelope Theorems for Non-Smooth and Non-Concave Optimization," Mimeo, University of Pennsylvania.

Fella, G. (2014): "A Generalized Endogenous Grid Method for Non-Smooth and Non-Concave Problems," *Review of Economic Dynamics*, 17, 329-344.

Lochner, L. (2011): "Chapter 2 - Nonproduction Benefits of Education: Crime, Health, and Good Citizenship," in *Handbook of The Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, vol. 4 of *Handbook of the Economics of Education*, pp. 183 – 282. Elsevier.