"If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is”.

~John Louis von Neumann
Preface

In 1968 Gary Becker published his article “crime and Punishment and economic approach”. This article could be seen as a new beginning for an approach to crime, long forgotten, in the history of philosophy and crime. It didn’t, as so many seems to believe, offer us a new theory of crime. What it did was to introduce a new methodology to an issue, which in more than 100 years had been dominated by a positivistic approach, which goal is to “cure the criminals from their sickness”. ¹ This research program is by nature extremely optimism, because it wants to detect and isolate all the relevant variables. The economic approach, starting with an axiom of “free will”, assuming people are solving a cost-benefit problem, is a much more generalized set-up, and therefore in reality, accepting a much lower ambitious level. Needless to say, positivist do find economist odd, generalizing and primitive, because the core in economics is “a ghost in the machine”, leading to a metaphysical concept of “a free will”. And of course, the concept of a free will, is in one respect absurd, because all human actions is in the end a consequence of social and genetic forces, meaning that all individual human actions, in theory, could be forecasted. However, what you need for this positivistic mission is access to all relevant data, a reasonable model and a very highly efficient computer².

In the science field of nature, the positivistic approach makes a lot of sense. But in the field of social science the main problem, is the degree of complexity and not at least, the problem, that a lot of relevant psychological data simply is not available in the statically sense. So, because of the lack of data, there is no way we can be sure to detect causality, forming a hypothesis and testing it by statically means, simply because statistics is only able to show us something about correlation. The positivistic approach means that you are choosing a path, where it is nearly impossible to detect which variable there is relevant, and which is not. That means that the evolution of science will come to a halt and the number of theories/papers simply will exploded, because there is no way to finding out who is right and who is wrong. So it should be of no surprises that the positivist approach to crime, had lead to an enormous amount of “relevant” or more precisely, speculative, theories, which, as a consequence, had lead policymakers in the dark.


² What we need is a daemon outside cosmos, as explained in causal or scientific determinism by Pierre-Simon Laplace in 1814.
Maybe the neoclassical approach to social science is not perfect. But the whole idea is based on that social science is very hard, which mean that we must somehow move back to some last resort of defense, a true axiom, from where we can build our theories on a logical foundation, using, where it is sensible, mathematical tools. That’s what Becker 1968 was all about, and that why we need to build theories of crime on a foundation on logic and choice. Taking the corn out of the positive approach will however mean that we should only be able to predict at the highest aggregate level. But this is, as I see it, where science can bring us.

This paper

Even though that the issue of crime, lately had more attention from economist, the attempt to building a microeconomic framework, regarding crime, seems too had been neglected. In this paper I will try to address that problem, and trying to take a look at, what I think is the central issues. I will show that microeconomic theory, starting by regarding people as optimizers allocating time, is actually fully compatible with many standard theories of crime, and that we therefore actually not will lose a lot of insight, moving forward with economic analysis. Especially, I will address the problem with discounting, which I believe, is sometimes confuses as irrationalistic behavior and being postulated as a very important factor in the phenomena we are calling crime. As I will explain, high degree of time discounting, could be seen as a highly specialized case, not the general case of crime.

However, trying to start with a metaphysical platform, individuals maximizing utility, with could be regarded within reason, as true, the scientific price paying for being right, is that you often will be to general and therefore lose predictable power. In the last section I therefore discuss the possibilities, that instead of analyzing crime in a consumer style framework; we could alternatively regard people committing crime, as single owned firm, which trying to maximize income. Of course, this will restrict economist from getting insight to all problems of crime, but actually preserved the basic issues. And I believe that most cases, crime is actually no more than a short cut to accumulated income, and punishment, a sort of tax on this kind of business. This means that we actually, more or less, are able to decide, how much crime we will tolerate.

---

1. **The individual allocation of time**

The basic workhorse in this paper is to try to take a slightly different route, or more precisely, a more generalized route, away from standard microeconomics. Instead of consuming goods, the agents are consuming time, $T$, on different sort of activities for example driven a car, eating hot dogs or stealing pancakes. Because human action always takes some time $t$, we define the Euclidian space as:

$$\mathbb{R}_+^n \equiv \{t_1 \ldots t_n | t_i \geq 0, i = 1 \ldots n \} \subset \mathbb{R}^n$$

We are using the standard assumption of rationality, which means that the agent’s preference must at least satisfy:

**Completeness:** for all $t_1, t_2 \ldots \in T$, we must have that $t_1 \succeq t_2$ or $t_2 \succeq t_1$

**Transitivity:** for all $t_1, t_2, t_3 \ldots \in T$, we must have that if $t_1 \succeq t_2, t_2 \succeq t_3$ then $t_1 \succeq t_3$

Of course it’s very easy to criticized these assumption, in special the assumption of transitivity (See for example; Kahneman and Tversky\(^4\) (1984)) However, because an assumption is not perfectly true, it’s not to say that such an assumption is not a very useful as a proxy for reality, moving forward. In the physical world, there is “no such thing” as a perfect vacuum or a frictionless plane. Making abstractions from “reality” frees us to focus on more important issues and allowing us to bring to bear powerful mathematics tools that help extend insight beyond the reach of intuition and experience. This is standard in every branch of science. Therefore we will simply ignore any critics of rationality and define a utility function $u(T)$.

**Definition:**
A function $u: t \rightarrow T$, is a utility function representing preference relation $\succeq$ if for all $t_1, t_2 \ldots \in T$,

$$t_1 \succeq t_2 \iff u(t_1) \geq u(t_2)$$

Completeness and Transitivity is some necessary condition for established a utility function. But simply for mathematical convenience, we also need the idea of continuity.

Definition:
The preference relation $\succ A$ is continuous if it preserved under limits. That is, for any sequence of pairs 
$\{(t^n_1, t^n_2)\}_{n=1}^{\infty}$, with $t^n_1 \succ t^n_2$ for all $n$, $t_1 = \lim_{n \to \infty} t^n_1$, $t_2 = \lim_{n \to \infty} t^n_2$, we have that $t_1 \succ t_2$.

Therefore there exist an continuous utility function $U(\bullet)$. For a simple proof see Colell\textsuperscript{5}.

If the assumption of Completeness, Transitivity and Continuity with respect to time consuming is satisfied we can then move to the question, how the rational agent will maximizing the utility, if standing in front of different possibilities of human action.

2.1 The solution to the standard problem – using time efficient

Notice that we at this point, say’s noting about prices at different sort of action. The only price for any action $t$, is simply the time involved. If $T \equiv 1$, we can write a time budget line as; $\sum_{i=1}^{n} t_i = 1$ which means that the agent facing the following problem:

\[
\begin{align*}
    \text{Max } & U(T) \\
    \text{St. } & \sum_{i=1}^{n} t_i = 1 \\
\end{align*}
\]

Because $t>0$, and $U(T)$ is continues, this maximizing problem have at least one unique solution if $U(T)$ is a monotonicity function.

Definition:
The preference relation $\succeq$ on $T$ is monotone if $t_1 \in T$ and $t_2 \succ t_1$ implies that $t_2 \succ t_1$.

Monotonicity is standard in microeconomics, simply says that “more is better”, meaning in this case that all human action is followed with some positive utility. Actually we don’t need this assumption to establish a unique solution for (1). The weaker assumption of locally nonsatiation is satisfactory. In any case the solution for this non linear programming problem is to form the Lagrangian;

\[
\begin{align*}
    \text{(2)}
\end{align*}
\]

\textsuperscript{5} Andreu Mas Colell., Michael D. Whinston and Jerry R. Green. Microeconomic theory, Oxford university press 1995
Following the Kuhn-tucker conditions:

\[
\frac{\partial L}{\partial t_i} = \frac{\partial u}{\partial t_i} - \lambda = 0, \quad i = 1 \ldots n
\]

Or if we let \( \nabla u(T) = \left[ \frac{\partial u}{\partial t_1}, \ldots, \frac{\partial u}{\partial t_n} \right] \) denote the gradient vector of \( u(\cdot) \) at \( t \), we can write in matrix notation

\[
\nabla^* u(T) = \lambda
\]

For any two set of human action \( n=2 \) we have:

\[
\frac{\partial u^*}{\partial t_k} = \frac{t_j}{t_k}
\]

Or put differently

\[
MRS_{t_j t_k} = \frac{t_j}{t_k}
\]

This says that the marginal rate of utility substitution between two sets of actions must equal the relative time involved in these actions. So the maximum principle in (5) simply says that a person will continue with some sort of action, until the marginal rate of substitution of those action is equal to the relative price (in this case time involved) between this two sets of action.

**An example, two actions and Cobb-Douglas utility**

Consider the case, where there exist only two different way for an person to allocate time \( (t_1, t_2) \). Our person facing a simple Cobb-Douglas utility function of the type \( U = t_1^{a} t_2^{1-a} \). The solution to the problem, following (2) is the solution to the problem, taking log:

\[
L = a \log(t_1) + (1-a) \log(t_2) + \lambda(1 - t_1 - t_2)
\]

The solution must solve the following system:

\[
\frac{\partial L}{\partial t_1} = \frac{a}{t_1} - \lambda = 0, \quad \frac{\partial L}{\partial t_2} = \frac{1-a}{t_2} - \lambda = 0, \quad \frac{\partial L}{\partial \lambda} = t_1 + t_2 - 1 = 0
\]
Because \( \frac{a}{t_1} - \lambda = \frac{1-a}{t_2} - \lambda \), and because \( t_1 + t_2 \equiv 1 \) it is straightforward to show that:

\[
(7) \quad t_1 = a \quad \text{and} \quad t_2 = 1 - a
\]

An geometrical interpretation for the par of \( t_1 \) and \( t_2 \) is:

Hence nothing in microeconomics in general says that we should exclude the possibility that \( a < 0 \), meaning that the activity \( t_2 \) is a bad. This will happen if we do not include the assumption of monotonicity. It’s typical more easy to simple included the assumption of monotonicity, meaning that \( a \geq 0 \). Hence we can interpret the parameter in this simple case, as the “strongness” not doing \( t_2 \). If \( a \to 1 \), then \( t_2 \to 1 \), then we have:

Hence nothing in microeconomics in general says that we should exclude the possibility that \( a < 0 \), meaning that the activity \( t_2 \) is a bad.
2.2 When some sort of human action generates punishment

Suppose we have a situation, where any human action \((t_1, t_2, \ldots, t_n)\) is accompanied with some action from a central planner, which for any sort of action could try/not try to imposed some sort of taxation or other form for punishment, \(\tau\), with that action. However, in the theory of crime, we normally accept, that the probability of punishment, \(\rho\), is not always equal to one, meaning that the relevant factor for the individual to consider is the excepted loss from doing crime, \(\phi\), which in this case could be considered as:

\[
\phi = \rho \tau
\]

So when \(\rho \to 1, \phi \to \tau\), meaning that the expected punishment must equal the actual punishment, when an individual is very sure to be detected\(^6\). It is important to think \(\phi\) as a vector, so that

\[
\phi = \{\phi_1, \phi_2, \ldots, \phi_n\}
\]

We will then say that:

\[
\frac{\partial u_i}{\partial \phi_i} < 0
\]

meaning that any central planner action following one action (called crime) from the individual, will be followed by a punishment, which will give the individual some negative utility impulse. We also recognize that \(\frac{\partial \phi_i}{\partial t_i} > 0\), meaning that the expected punishment must be a increasing function of the activity level in \(t_i\).

Therefore the individual now must try to maximize the following system;

\[
\begin{align*}
\text{Max } U &= U(t_1, t_2, \ldots, t_n, \phi_1(t_1), \phi_2(t_2), \ldots, \phi_n(t_n)) \\
\text{St. } \sum_{i=1}^{n} t_i &= 1
\end{align*}
\]

If we simply defined the net benefit from doing any action \(t_i\) as

\[
U(B_i) = U(t_i) - U(\phi(t_i))
\]

\(^6\) Hence we are following the first rule of Bentham (1931) as; “The evil of the punishment must be made exceed the advantage of the offense” And later in chapter 2, Bentham included probability in his second rule as; “the more deficient in certainty a punishment is, the severer it should be”.
The system simply collapse to

\[ (10) \quad \max U(B) \]
\[ \text{St.} \sum_{i=1}^{n} t_i = 1 \]

With the solution as in (4) or reformulated as:

\[ \frac{\partial u(B_k)^*}{\partial t_k} * t_k = t_j * \frac{\partial u(B_j)^*}{\partial t_j} \]

Geometrically we can interpret this as;

![Graph showing the relationship between utility and time allocation with and without punishment.]

So a higher price of some action \( t_i \) will not surprisingly mean lower supply of that action. Alternatively we could interpret this as,
Expected Cost of action $i$

We could define the elasticity of the supply, with respect to $t_i$, as:

$$\varepsilon_i = \frac{\partial t_i}{\partial \phi_i} \cdot \frac{\phi_i}{t_i}$$

Or because $\phi = \rho \tau$

We get:

$$\varepsilon_i = \frac{\partial t_i}{\partial \rho \partial \tau} \cdot \frac{\rho \tau t_i}{t_i}$$

Which of course must be a negative number, because we expect a negative utility from punishment? Hence the conclusion is straightforward. The fundamental logic of microeconomics, would without any discussion, simply implied, that when there exist a human action a central planner dislike, and if that central planner imposed a punishment (price) for that action, as the individual dislike, then the individual will choose, all other thing being equal, to supply less of that action! How much the human action responds to $\phi$ is of course an empirical question, but as I see it, NOT necessary the most relevant question!

Summary we get: The supply of any human action, depends on:

1) The preference or utility for that action
2) The benefit from alternative action
3) The expected loss of utility by punishment from that action. Either imposed as a rising probability from detection $\rho$ or a rising punishment, $\tau$ when being detected

So the if the standard assumption of microeconomic theory is fulfilled, then we can concluded that the generalized theory of crime, seems to imply that the characteristics of a criminal seems to be a person, which don’t dislike crime behavior very much, do not have a lot of very good time alternatives, and seems NOT to be very deterrent from punishment. The last question seems particularly interesting for economist, because it seems plausible to maintain that one scheme of crime is that it generating utility today, but the

---

7 As I see it, it seems much more relevant to figure out the “correct” price of any action. The “correct” price of crime could be seen as the society’s external cost, as the polluter (criminal) must pay back.
cost come (maybe?!) tomorrow. What you are doing today must be paid with some stream of cost of utility in the future, meaning it’s not really the utility value of action we are considering, but the discounted utility action of that action. Second generation of models, for example (McCrary, Lee 2005), (Imai 2004), (Lochner 2004) (Fent,Zaleak,Feichtinger 2009) seems to interpret this discounted utility scheme.

2.6 The discounting of punishment (the special case)

At this stage, lets us consider some rather simple generalized function, determined the agent’s net present utility of crime in the period between: [0, a]. Suppose that the agent discount factor could be handle as a constant, which we will denote \( \theta \).

The generalized maximization problem then becomes:

\[
\begin{align*}
\text{Max } U(t) &= \int_{0}^{a} U(t_1, t_2, \ldots, t_n, \phi_1(t_1), \phi_2(t_2) \ldots \phi_n(t_n)) e^{-\theta t} dt \\
\text{St. } \sum_{i=1}^{n} t_i &= 1
\end{align*}
\]

The solution in this problem is “not dynamic” because the agent simply maximize at the particular point in time. Therefore the solution is:

\[
\frac{\partial u(B_k)}{\partial t_k} * t_k = t_j * \frac{\partial u(B_j)}{\partial t_j}
\]

An therefore exactly the solution as in (8) or (4). Hence nothing has been gained in the analysis, just simply to include the discount factor in the problem. The discount factor just means that the present value of punishment will somehow be diluted. To see this more easily, let’s define: \( t_i \) as all possible human action. \( t_j \) as all possible sort of non criminal act. And \( t_k \) as all possible sort of criminal acts, which generate some sort of expected punishment \( \phi \).

\[8\] In the book “a general theory of crime 1990” two sociologist Michael R. Gottfredson and Travis Hirschi, seems to verbally consider the same idea.
So $t_i = t_j + t_k$ and $t_j \in t_i$ and $t_k \in t_i$. If a person picks his action purely from elements in $t_j$ in the time interval $[0, b]$ we have that $t_i = t_j$, and the net present discount value of that sequence of action (npv) could be determined as:

$$npv(t_j) = \int_{0}^{b} e^{-\theta T} u(t_j) dT$$

If a person picks his action from all elements in $t_i$, the present discount value must take into account, that there is some expected punishment, which we will suggest will take place. That means the agent facing the following utility maximizing problem of time allocation between 0 and a:

$$npv(t_i) = \int_{0}^{b} e^{-\theta T} u(t_i) dT + \int_{a}^{b} e^{-\theta T} u(t_i) dT$$

Hence the criminal must take into account that in the time interval $[0, b]$ there will be some punishment from the period $[a, b]$:

Hence a criminal person must take into account the discounted net loss (The area from a to b), when he decide some illegal action in the time interval $[0, a]$ which give some benefit (the area from 0 to a). It is evident, that,

$$\lim_{\theta \rightarrow \infty} \rho(t^*)u(c(t)) \rightarrow 0$$

So as $\theta$ grows larger, then the expected loss from crime becomes smaller! Hence there can be no deterrence of a person, which have a very large discount factor, because he simply act impulsive (not thinking about the future very much). Notice however, that a high discount factor does not imply anything
regarding the analysis of rational/ irrational behavior. A high discount factor could easily be interpreted as highly rational. What does mean is however, is that a central planner is not able to fully control crime by deterrence, if some persons have a very high discount factor. In such a case the elasticity is 0.

How likely is this situation? The answer is that it is not very likely. Some people truly are insane, and therefore, seems to have and extremely high discount factor. And some people truly are fanatics, totally not afraid for any sort of deterrence. But this must be special stories, maybe interesting in the news, but absolute not the standard case, regarding crime. Think about it. Even an animal is able to discounting. Thief does indeed go out of night. People do not sell drugs at Sachs of 5.avenue. And nobody seems to be afraid being in the middle of day in central park. What people seems too worried about is situations where the is small changes that bad persons being caught, for example in “bad neighborhoods, when it is dark”, which is just another way of saying, that we actually believe that in general, the discount factor is not infinite.

That’s not to say, that the discount problem shouldn’t be a consideration. I believe it should. It’s actually an important issue, even if the discount is “normal”, because typically you are trying to punish people with time, sending them to jail. Even a small discount factor could easily dilute the deterrence of time spent in jail, meaning that jailing is not always a very effective way controlling crime. Especially not, if you’re at the same time had a very low discounted income. This is just another way of saying, that fine or directly monetary compensation to the victims, could be considered, in many cases, as a more effective and efficient way of raising the cost of doing crime⁹.

The Income approach

The dilution of punishment by time is of course an important question, but as such, it has NO consequences for the overall and general maximizing problem for the individual. For analytical purpose, nothing had been gained to consider the discount factor, with could be consider as a problem as it own. And as we also have seen, even to include a high discount factor, is not a very interesting case, at least not for an economic point of view, because it just means, that people do not respond accordingly to prices. So this, as I see it, leaves us with the question, why we should included the utility scheme, regarding crime, in the first place? Maybe, Bentham was wrong, trying to analyze, crime as utility cost-benefit problem, because this surely

---

⁹ This is the main point in Becker 1968, and as I see it, one of the most important points in economics regarding crime. Why do we calculate the price of human action in time (time spent in jail) instead of calculating the price in monetary terms, as we actually did before “the birth of the prison” (see for example Foucault, Michel. Discipline and Punish: The Birth of the Prison 1975)
will catch the general idea, but also leave us with such a general problem, that we should not be able to logically predict very much. Why not simply said, that individuals committing crime, should be regarded as one-man companies, trying to maximize income? Indeed. 98% of all crimes are directly related to the problem of transfer of property and lots of the last 2% is directly or indirectly related to that issue as well, for example as a defense for a drug business territorial. Hence, skipping people which is temporally or permanently insane (a high discount rate), we should be able to present the individual one-man firm, as trying to maximize his income, choosing the best way to produce income as:

\[
\text{max } l = W_1(t) \cdot t_1 + W_2(t^*) \cdot t^* - \rho \tau(t^*)
\]

s.t. \( t + t^* = 1 \)

Where time allocated to crime is denoted \( t^* \) and where the offsetting price for that action called crime is denoted \( \tau \) there is coming with a probability of \( \rho \). Because you spent your working time at one work, you cannot have an income of the other, so the opportunity cost is \( W_1(t) \cdot t_1 \). Lets assume that the wages for a legal job is a constant, meaning income from “normal” work is \( W_1 \cdot t_1 \) and also, for simplicity assume that expected loss, doing crime, is linear, \( \rho \tau = \phi \). It should be quite resonable to assume that the agent is facing decreasing return from doing crime, because “the low fruit” from crime projects, will be done first, meaning that: \( \lim_{t^* \to \infty} W_2(t^*) \rightarrow 0 \).

Of course, this idea is not very elegant compare to the concept of utility maximizing, but the result is straightforward and intuitive. Forming the lagrange function:

\[
L = w_1 \cdot t_1 + W_2(t^*) \cdot t^* - \rho \tau(t^*) + \lambda(1 - t_1 - t^*) = 0
\]

And we get:

\[
\frac{\partial l}{\partial t^*} = -w_1 + W_2'(t^*) - \phi = 0
\]

Or:

\[
w_2'(t^*) = \phi + w_1
\]
Which is exactly as we should expect, because the marginal benefit from doing crime, must in an optima, be equal to the cost of doing crime, which is the loss from others types of work and the expected cost from doing crime. Graphically this says that:

![Graph showing the relationship between income from crime and time allocated to crime.](image)

(Point a is the optimizing amount of time, allocated to doing crime, and b is the optimizing amount, after a punishment had been introduces)

As we should see, it is nearly impossible to eliminate all crime-projects, because \( \lim_{t \to 0} W_2(t^*) \to \infty \). However, it is easy to spot, that a higher expected price, doing crime, must lower the relative attractiveness, so crime activity must fall. There is simply no way around this problem. Hence, simple basis economics, is telling us that crime depend on:

- Probability being caught,
- the punishment when being caught,
- the opportunity cost (going wages rate)
- and the value of crime.

**Conclusion**

Adding the utility approach and the income approach together, we should be able to give some sensible economics solution to the analytical problem of crime. Starting by the axiom of maximizing, I think, it is easy to see that this is fully compatible with, what most people actually think is the core of crime? Economic theory gives us a logical starting point, and we are of course capable to include more variables, as example, human capital, and let the analytical framework becomes even more realistic. As I see it, ripping
the problem apart, building it up block by block, testing it piece by piece, is the right way to go, and the way which can lead us to a better understanding of the logic and incentives of the world of crime. However, a better understanding of incentives is only a part of the more complex story and the ultimate goal of economic theory, which is how to minimize the welfare loss of crime. I am sure, that the overall conclusion, should not depart very much, from the general conclusion from the theory of externalities, that minimizing the cost of crime, means that price of doing crime, should be no more and no less, the welfare loss, from that action.