Measuring offender discount rates

Andrew Torre[1], Darren Wraith[2],

1 Deakin University 221 Burwood Highway. Burwood Victoria 3125. andt@deakin.edu.au
2 Quennsland University of Technology.. Gardens Point. GPO Box 2434 Brisbane Queensland. 4001
d.wraith@qut.edu.au

It is commonly asserted that one characteristic of people who engage in crime is that they have high discount rates at the time of committing the offence. While discount rates have been inferred for a wide variety of decisions in different contexts, there is an absence in the literature of empirical estimates for offenders. In this study, the authors attempt such an exercise through an examination of the plea decision of a sample of individuals prosecuted for murder, aggravated robbery and theft in the NSW higher courts in Australia in 2004.

1

1. Introduction

According to Becker (1968), a rational individual will compare the expected return from illegitimate, with the return from legitimate activity when deciding how much time to allocate to each one. Alternatively, possibilities such as adverse social conditions, poor upbringing, mediocre school performance, lack of parental discipline, genetics and psychological problems (Buchanan & Hartley, 1992) lower the return from legitimate relative to illegitimate activity, predisposing individuals to the latter. Computing the expected return from crime involves an inter temporal comparison since the benefits are immediate, however the prospective expected cost depends on the individual's probability of getting caught and convicted, and the ensuing penalty. Consequently, the weight the individual places on these two magnitudes will reflect his or her discount rate.

Positive time preference is commonly assumed in the economic literature and its theoretical foundations are well established (Olson and Bailey 1981). Individuals, who place a much higher value on present relative to future returns from an activity than society as a whole, will have a relatively high discount rate, and those who do the opposite will have a negative discount rate. Many writers suggest that at the time of completing criminal acts, individuals have high to very high positive rates of time preference. For example, Gottfredson and Hirschi (1990) argue that the criminal act is governed primarily by short- term pleasures, and only secondarily if at all by the threat of long term pains. People who commit a crime all tend to choose short-term advantages over long term costs, and this tendency declines with age. Wilson and Hernstein (1985) focus on time discounting in explaining criminal behaviour, arguing that since the rewards usually precede the costs of crime, the impact of a tendency on the part of people to discount future benefits and costs is to consistently increase the likelihood that crimes will be committed. As well as important implications for the crime rate, high rates of time preference imply that sentence length has a reduced impact on crime.

There have been several empirical studies on the extent to which people devalue future rewards, and these have elicited evidence of high to very high discount rates in the context of different decisions. Economic decisions from which discount rates have been inferred include purchases of consumer durables, saving and inter temporal labour-leisure substitution (Loewenstein & Prelec, 1992). One example is Hausman (1979) who undertook a study of actual air-conditioner purchases across different income groups, and showed that in accepting higher operating costs in return for lower purchase prices, consumers devalued the future at annual rates as high as 89% with the mean rate being 25% (Ainslie, 1992). In another group of studies reported in Ainslie, people were asked how they would trade off the amount and the delay of extra income that was entirely hypothetical. Implicit discount rates varied from 36% to 122% (Kurz, Spiegelman, & West, 1973); 60% (Benzion et al., 1989); 100% to 120% (Maital & Maital, 1977) and 5,000% (Lea, 1977).

Empirical studies and estimates of offenders' discount rates are virtually non-existent in the large time preference literature, and although individual discount rates are not directly observable, they should ideally be inferred from individual decisions. In this paper, the authors attempt an estimate of discount rates for a sample of offenders processed in the criminal justice system of NSW in Australia in 2004. The study is different from the usual methodology for estimating rates of time preference. It is not conducted experimentally by, for example, asking offenders hypothetical questions about tradeoffs between illegal returns for different time periods and probabilities of apprehension and conviction. A straight forward calibration, rather than estimation exercise is carried out utilising actual decisions; not the crime decision itself, but the subsequent plea choice of those offenders who are caught and prosecuted. As with the former, the outcome of the latter decision is temporally remote, since inevitably defendants will confront court delays. Consequently, the defendant is forced to make a comparison between immediate and delayed consequences. Only not on bail defendants following a guilty plea or trial are examined since they invariably receive the harshest form of punishment upon conviction, a jail sentence backdated to when they were charged and remanded in custody. First, the theoretical model underpinning this study is briefly set out, second the empirical methodology is described, and third the estimates and a brief discussion are presented.

2. The Theoretical Model

We assume that the defendant is not on bail and legally aided, which enables us to write the cost identity in its simplest form, i.e. the cost of the guilty plea or prosecutor's final offer is equal to:

$$C = (Y, D, r, t) \tag{1}$$

where Y is the defendant's foregone income while on remand waiting for the court hearing until the final disposition of the case, and then during the duration of the sentence D, r is the offender's discount rate and t is the time elapsing from the time of remand until the final court hearing at which the defendant is sentenced.

The present value of expression (1) is given by (2).

$$C = Y\int_0^t e^{-rt}\,dt + Y\int_t^{t+D} e^{-rt}\,dt - Y\int_0^t e^{-rt}\,dt$$

$$= \frac{Y}{r}(1-e^{-rt}) + \frac{Y}{r}(e^{-rt} - e^{-r(t+D)}) - \frac{Y}{r}(1-e^{-rt}) \qquad (2)$$

If the not on bail defendant chooses to plead not guilty, the expected cost identity of a trial is equal to:

$$E(C) = (Y, P, D, r, T) \qquad (3)$$

where P is the probability of being convicted following a trial, and T is the time elapsing from the time of remand until the conclusion of the trial.

The present value expression of (3) is given by (4).

$$E(C) = Y\int_0^T e^{-rT} dT + [(P\int_T^{T+D} Y e^{-rT} dT - (1-P)\int_T^{T+D} Y e^{-rT} dT) - (P\int_0^T Y e^{-rT}\ dT$$
$$- (1-P)\int_0^T Y e^{-rT}\ dT)\ ]$$

$$= \frac{Y}{r}(1-e^{-rT}) + (2P-1)[\frac{Y}{r}e^{-rT} - \frac{Y}{r}e^{-r(T+D)}) - \frac{Y}{r}(1-e^{-rT})] \qquad (4)$$

It is important to note that (3) and (4) are identities not cost functions and therefore not a hypothesis about the drivers of the each plea, consequently, the statements are merely definitional, i.e. the cost and expected cost respectively are defined as the defendant's discounted foregone and expected foregone income during incarceration. Also, we do not attempt to posit a relationship between P and time, since we do not know what P does over time. In some cases it will increase, and in others it will decrease, so that it moves randomly over time. Since there is unlikely to be a deterministic relationship between P and time, in our simulation exercise we use a range of constant P values to gauge the sensitivity of our results to these different values.

3. Methodology

The defendant will choose the lower cost option out of (2) and (4) and this will change as court delay varies. When formulating the plea decision, each defendant will confront a large number of plausible combinations of each of the parameter values in (2) and (4). For each combination of Y, t, T, D + t, D + T and P, our methodology is to find values of r, which we denote as r*, that equate the cost of a guilty plea or prosecutor's final offer with the expected cost of a trial in (2) and (4). The relative attractiveness of the expected values of the two courses of action is related monotonically to the discount rate, and as a result, the observed pattern of plea bargains, when combined with values of the other key variables yields a set of values for the discount rate r*, that partitions the plea space into that where the guilty plea is preferred and conversely. As an illustrative example, a defendant confronts the likely following parameter values ex ante; Y = $52,292; T = 1.03 years for a trial and t = 0.8136 years for a guilty plea; D = 0.5 years for a not guilty plea and 3 years for a guilty plea and P = 0.3.

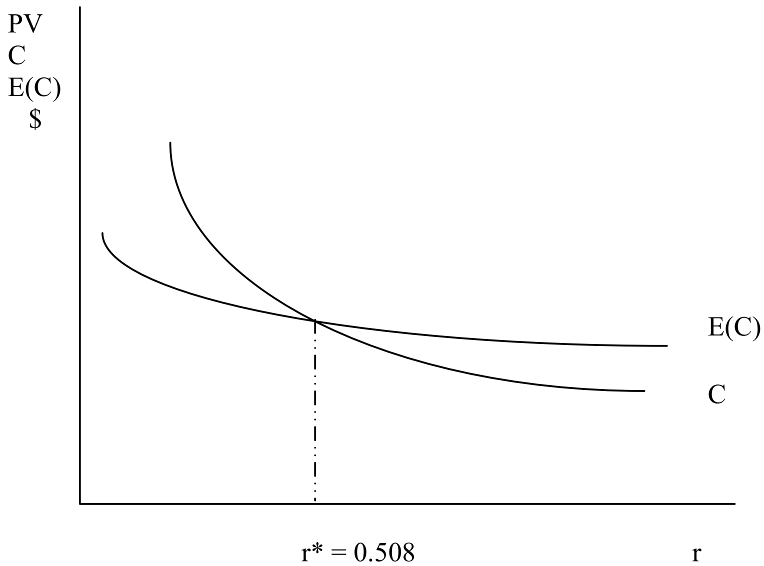


Figure 1

In figure 1, at r* = 0.508 the defendant is indifferent between pleading guilty and going to trial since the cost of the guilty plea or certain prison sentence C equals the expected cost of the trial E(C), or the expected prison sentence. The decision rule is to go to trial if $r < r^*$ since $E(C) < C$ and plead guilty if $r > r^*$ since $C < E(C)$, where r is the defendant's actual discount rate. Consequently, if the defendant in figure 1 goes to trial, his or her unobservable discount rate must have been ≤ 0.508, and conversely if the guilty plea is chosen, it must have been ≥ 0.508.

Alternatively, some parameter combinations will yield curves where at $r < r^*$, $C < E(C)$ and conversely at $r > r^*$ as in figure 2 below.

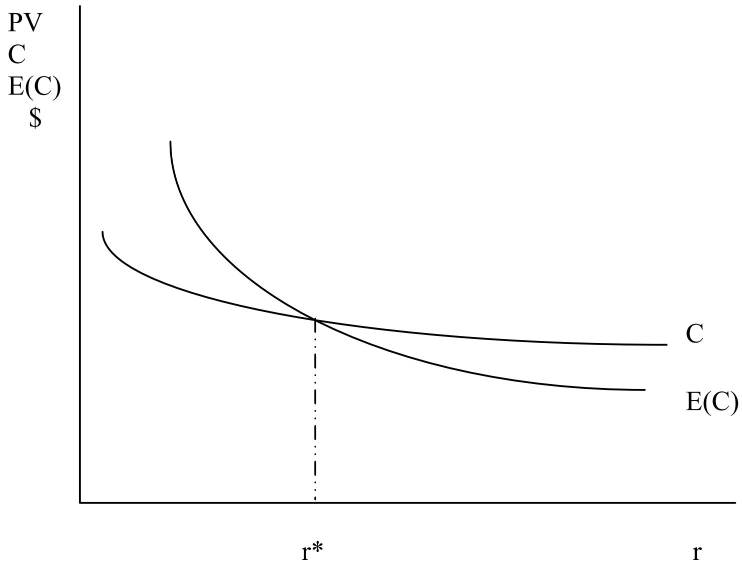


Figure 2

In this instance, if the defendant elected a trial, his or her discount rate must have been ≥ r* and if the guilty plea were chosen, then r* would be the upper bound estimate of the true r. Substituting all of the possible independent combinations of income, waiting times, expected sentences and conviction probabilities into (2) and (4) that a defendant would confront will generate a large distribution of values of r*, two such r*s are identified in figures 1 and 2.

We attempt to identify the entire distribution of r*s for three different offences in NSW in 2004, murder, aggravated robbery and theft, which were finalised in the NSW higher courts. Table 1 provides summary statistics for actual waiting time from remand until sentence is passed by the court defined in months, and sentence length (imprisonment in months) obtained from the NSW Bureau of Crime Statistics and Research (2004).

Table 1: Summary statistics for actual waiting time and sentence length

	Aggravated Robbery		Murder		Theft*
	NGP	GP	NGP	GP	GP

Waiting time (months)	Min	4.47	2.30	7.07	9.00	4.29
	Q1	9.73	6.75	18.57	14.52	6.87
	Median	13.00	9.43	22.53	16.94	8.26
	Mean	15.99	11.59	23.98	24.23	11.11
	Q3	18.10	14.52	26.82	22.00	15.58
	Max	39.77	59.50	85.94	80.26	21.06
	N**	53	240	66	15	39
Sentence Length (months)	Min	18	4	17	18	2
	Q1	39	18	156	114	6
	Median	48	24	174	171	8.50
	Mean	51.05	30.26	210.10	204.60	11.97
	Q3	60.00	39.00	198.00	198.00	14.25
	Max	90.00	93.00	999.00	999.00	36.00
	N**	19	240	21	15	39

Note: * there were only 3 Not Guilty cases.

N** is the number of not guilty (NGP) and guilty pleas GP.

For the computations, waiting time, sentence length and income estimates are based on a random sample of 19, 15 and 30 values for aggravated robbery, murder and theft respectively. Actual values are used for waiting time and sentence length, and for the unknown variable Y, we assume that income is normally distributed with a mean of \$50,000 and a variance of \$100,000, (standard deviation \$10,000 and range \$20,000 to \$80,000). It should be noted that the choice of income will have no impact on the position of the curves in figures 1 and 2 because a change in Y will shift both curves by the same proportion so that the value of r* will not change. For P the other unknown, we use three values 0.3, 0.5, and 0.8. In the case of theft, since there were only three not guilty pleas, we assume that the waiting time for a trial is 1.5 times the wait for a guilty plea, after inspection of waiting times for other offences as shown in Table 4. The sensitivity of this assumption is tested separately in Table 5 by assuming the waiting time for a NGp is only 1.25 times that of a guilty plea. Defendants prosecuted for aggravated robbery therefore faced 19^3 and 19^4 possible combinations of values of D, t, and Y for a guilty plea and D, T, Y and P for a not guilty plea disposition respectively. In the case of murder the maximum number of feasible combinations was 15^3 and 15^4 for the guilty and not guilty plea respectively, and for theft 30^3 and 30^4.

For each combination of values for the variables, we searched for an intersection point between the cost of a guilty plea and expected cost of a trial. These values of r* generate a distribution from which it is possible to infer values of individual defendant's actual discount rates, which are available from the authors. The results of this exercise are presented in Tables 2 to 5 below

Table 2: Results for Aggravated Robbery

Measures	P = 0.3	P = 0.5	P = 0.8	
	NG < GP	NG < GP	NG < GP	NG > GP
	r*	r*	r*	r*
Min	0.03	0.03	0.25	0.03
Q1	0.44	0.41	0.53	0.09
Median	**0.59**	**0.57**	**0.92**	**0.21**
Mean	0.62	0.63	1.15	0.25
Q3	0.77	0.85	1.62	0.29
Max	1.55	1.69	2.76	0.91
95% CI	(0.09, 1.35)	(0.09, 1.39)	(0.27, 2.68)	(0.03, 0.75)
N**	6460	5795	4028	1748

N ** Number of intersection points or values of r* found.

Table 3: Results for Murder

Measures	P = 0.3 NG < GP r*	P = 0.5 NG < GP r*	P = 0.8 NG < GP r*	 NG > GP r*
Min	0.13	0.05	0.13	0.03
Q1	0.31	0.31	0.37	0.04
Median	**0.39**	**0.41**	**0.51**	**0.09**
Mean	0.38	0.39	0.52	0.09
Q3	0.45	0.49	0.67	0.15
Max	0.69	0.75	1.17	0.17
95% CI	(0.15, 0.61)	(0.09, 0.63)	(0.13, 1.15)	(0.03, 0.17)
N**	3375	3195	2925	900

N** Number of intersection points or values of r* found.

Table 4: Results for Theft: Waiting time for a NGp is 1.5 times Gp

Summary measures	P = 0.3	P = 0.5	P = 0.8	
	NG < GP	NG < GP	NG < GP	NG > GP
	r*	r*	r*	r*
Min	0.03	0.03	0.03	0.57
Q1	0.17	0.19	0.77	0.57
Median	**0.35**	**0.39**	**1.09**	**0.58**
Mean	0.39	0.43	1.02	0.58
Q3	0.53	0.65	1.25	0.59
Max	1.19	1.19	2.08	0.61
95% CI	(0.05, 0.99)	(0.07, 1.05)	(0.03, 2.08)	(0.57, 0.61)
N**	7440	6210	6840	900

N ** Number of intersection points or values of r* found.

Table 5: Results for Theft: Waiting time for a NGp is 1.25 times a Gp

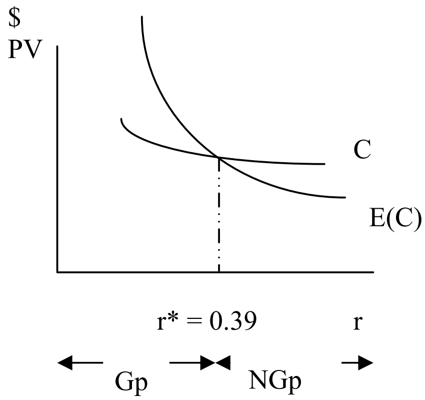
Measures	P = 0.3 NGP < GP r*	P = 0.5 NGP < GP R*	P = 0.8 NGP < GP r*	 NGP > GP r*
Min	0.03	0.03	0.03	0.11
Q1	0.19	0.25	0.45	0.11
Median	**0.39**	**0.47**	**1.09**	**0.32**
Mean	0.42	0.52	1.03	0.34
Q3	0.61	0.73	1.69	0.59
Max	1.39	1.61	2.34	0.59
95% CI	(0.05, 1.03)	(0.05, 1.25)	(0.05, 2.30)	(0.11, 0.59)
N**	10950	9150	10530	480

N ** Number of intersection points or values of r* found.

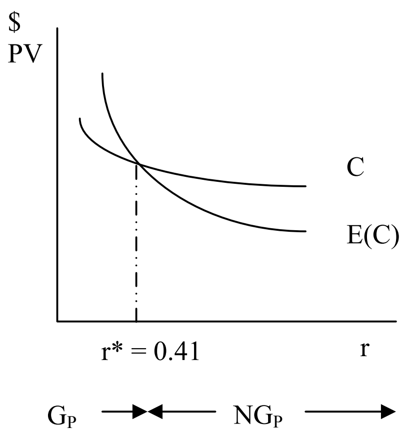
The results in tables 2 to 5 are summarised diagrammatically below. The interpretation follows in section 4.

Murder (median estimate):

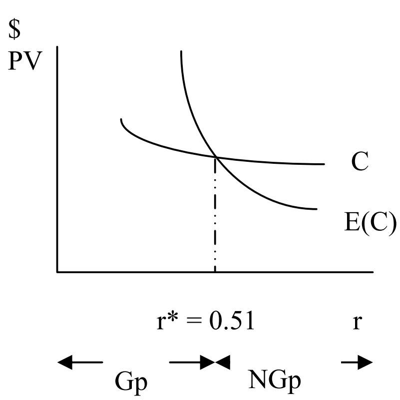
P = 0.3



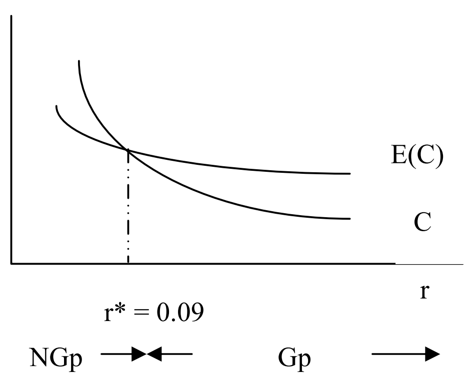
P = 0.5



P = 0.8

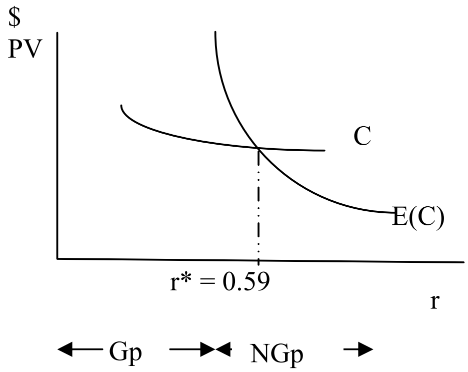


P = 0.8

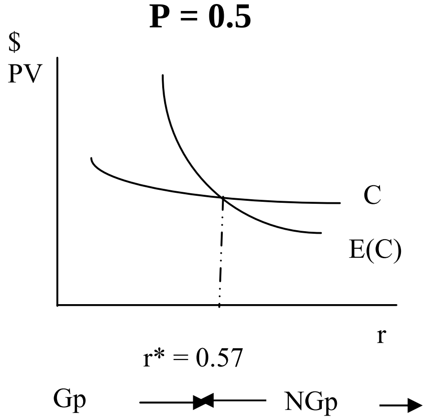


Aggravated Robbery (median estimate):

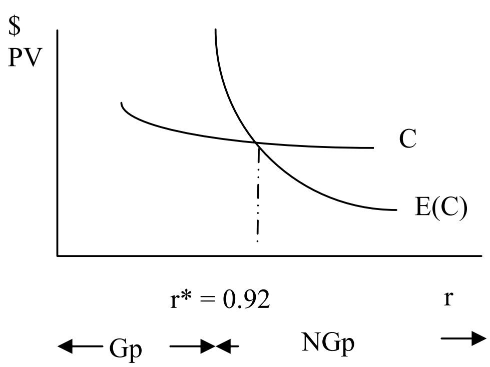
P = 0.3



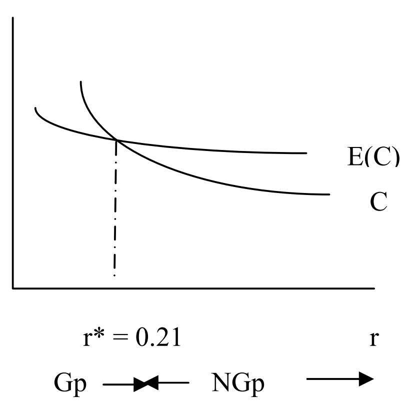
P = 0.5



P = 0.8



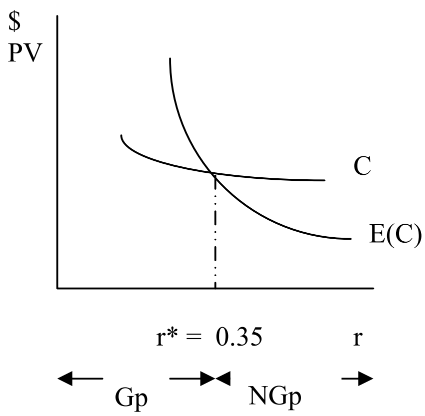
P = 0.8



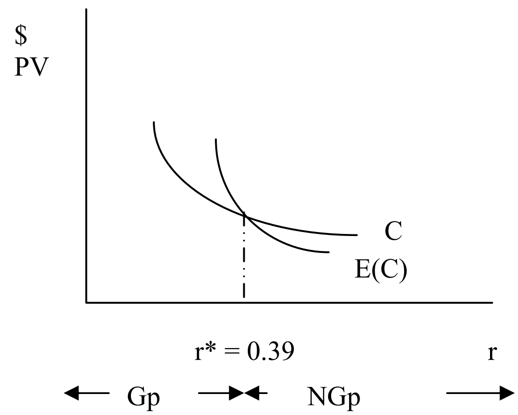
Theft: Waiting time for NGP is 1.5 times GP **1.25 times GP**

Median Estimates.

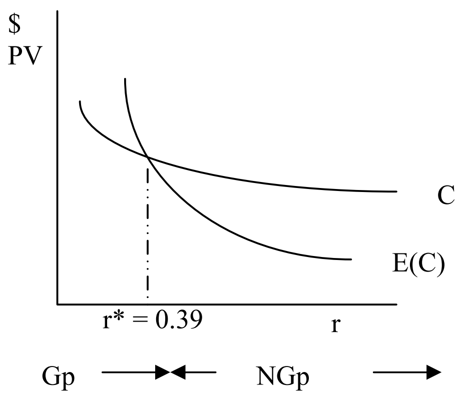
P = 0.3



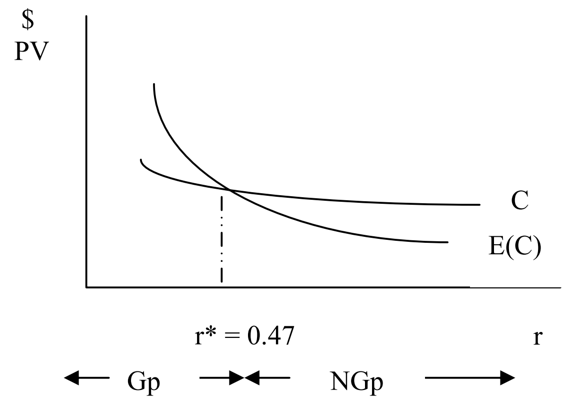
P = 0.3



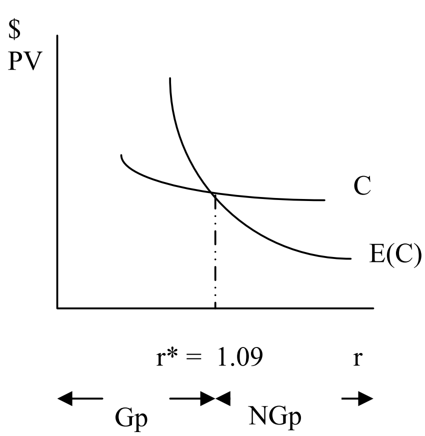
P = 0.5



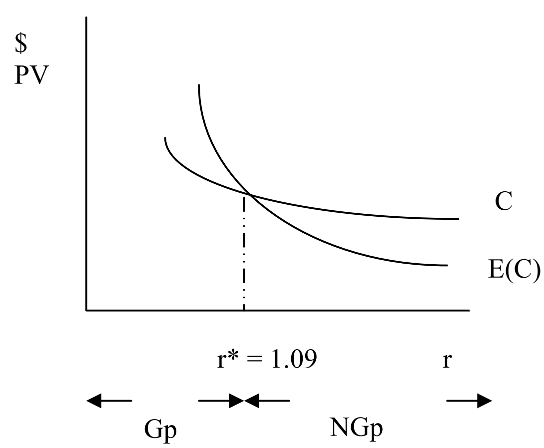
P = 0.5



P = 0.8

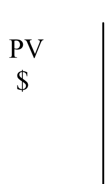


P = 0.8



P = 0.8

P = 0.8



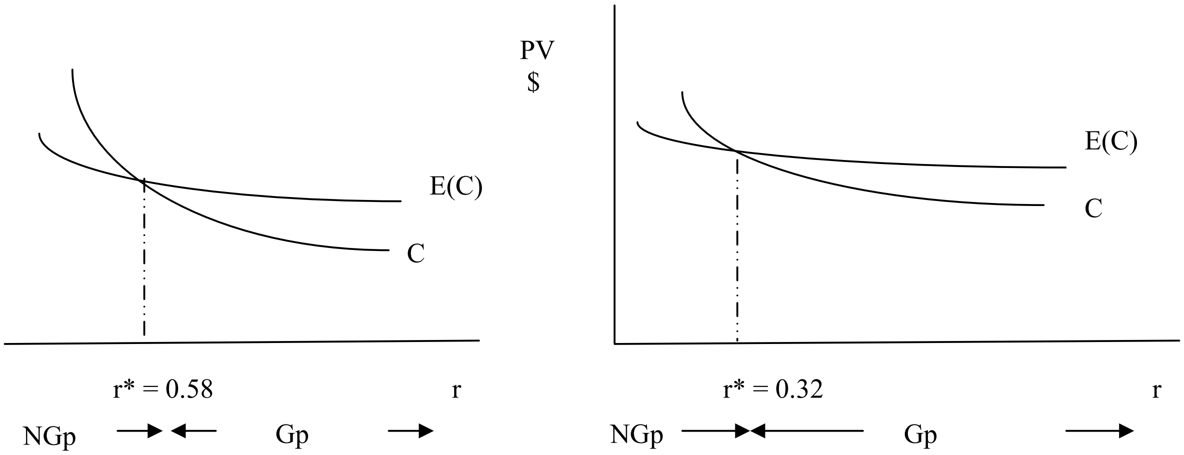


Figure 2: Results of the simulation exercise

4. Discussion

Given the distribution of the underlying data from which the samples were taken and since the r* estimates are on a continuous scale, the median is the most likely estimate of r* for each offence. Consequently, we compared the median values for the three offences to see if they were different from one another using Mood's Median Test, a non parametric test, which is a more robust alternative to the Kruskal-Wallis test in the presence of outliers in the data (Hollander & Wolfe, 1973). The results for all three offences were found to be significantly different ($p = 0.05$) for all three P values.

Since the payoffs are foregone income or costs, the inferred discount rates have a neat interpretation, the willingness to pay to defer the monetary cost of punishment by a month/$ of income. Table 6 highlights the willingness to pay estimates to defer certain punishment per dollar of income following a guilty plea.

Table 6: Summary: Willingness to pay to defer the monetary cost of certain punishment per dollar of income by a month.

Offence.	**Guilty Plea.**

Murder.	0.09 to 0.51c/dollar of income.
Aggravated Robbery	0.21 to 0.92c/dollar of income.
Theft	0.58 to 1.09c/dollar of income.

From table 1, the median sentence in years for murder, aggravated robbery and theft was 14, 2 and 0.85 years respectively, yet from table 6, the maximum willingness to pay to delay foregone income appears to be negatively and not positively correlated with the severity of punishment, as it is the lowest for murder even though the penalty was substantially higher than for the other two offences.

In effect, our estimates can be interpreted as implicit premiums over and above the unskilled wage rate for each of the three offences, and they provide further evidence of very high returns to some criminal offences against property. For example, according to Stevenson et. al. (2001) the estimated median value of weekly earnings for burglars in NSW is $2,000 yielding an annual income of $104,000. Recently, quite a lot of work has been done on deducing parameters from observed decisions using dynamic programming models, and Wolpin (1996) provides a good introduction to this literature. While a dynamic programming framework to the problem has not been utilised in this paper, the concept of bringing together actual decisions and a constructed model to explain these to infer an unobservable parameter using an alternative framework is the same. The estimation of returns from different offences and why they persist is an important avenue of future research into the formulation of crime reduction policies.

References

Ainslie, G (1992) *Picoeconomics. The Strategic Interaction of Successive Motivational States within the Individual.* Cambridge University Press.

Becker, G (1968), “Crime and Punishment: An Economic Approach”, *Journal of Political Economy*, Vol 76,, pp 169-217.

Benzion, U, Rapoport, A and J. Yagil (1989), “Discount Rates inferred from Decisions: An Experimental Study,” *Management Science,* 35. pp 270-284.

Buchanan, C and P. Hartley (1992), *The Economic Theory of Crime and its Implications for Crime Control*, The Centre for Independent Studies. NSW.

Gottfredson, M and T. Hirschi (1990), *A General Theory of Crime.* Stanford University Press, Stanford, California.

Hausman, J (1979) “Individual Discount Rates and the Purchase and Utilization of Energy Using Durables,” *Bell Journal of Economics,* 10, pp 33-54.

Hollander, M and D. Wolfe (1973) *Nonparametric Statistical Methods*, John Wiley & Sons.

Kurz, M, Spiegelman, R and R. West (1973) *The Experimental Horizon and the Rate of Time Preference for the Seattle and Denver Income Maintenance Experiments: A Preliminary Study.* Research Memorandum 21. Menlo Park, CA: SRI International.

Lea, S (1977) Delayed gratification in normal adults: A questionnaire study. Unpublished manuscript. Cited in Ainslie, G *Picoeconomics.* C.U.P.

Lowenstein, G and G. Prelec (1992) “Anomalies in Inter temporal Choice, Evidence and an Interpretation,” *QJE,* 107:2, pp 573-597.

Maital, S and S. Maital (1977) Time preference, delay of gratification and the intergenerational transmission of economic inequality: a behavioural theory of income distribution. In O.C Aschenfelter & W. E. Oates (Eds), *Essays in labor market analysis.* New York: Wiley.

NSW Bureau of Crime Statistics and Research (2004) NSW Criminal Courts Statistics.

Olson, M and M. Bailey (1981), “Positive Time Preference,” *Journal of Political Economy,* Vol. 89, No. 1, pp 1-26

Stevenson, R, Forsythe, L & D.Weatherburn (2001) “The Stolen Goods Market in New South Wales, Australia. An Analysis of Disposal Avenues and Tactics.” *British Journal of Criminology.* Vol. 41. Pp.101-118.

Wilson, J and R. Hernstein (1985), *Crime and Human Nature.* Simon and Schuster, New York.

Wolpin, K (1996) “Public Policy uses of Discrete–Choice Dynamic Programming Models,” *American Economic Review,* Vol. 86, pp 427-432.