5

Is Age the Primary Influence on Offending?

Orientation

In Chapter 2, we proposed a theory based on the idea that convictions occurred at random (according to a Poisson process). We proposed that there were three categories of offenders: high-risk/high-rate, high-risk/low-rate, and low-risk/low-rate. Both the risk of reconviction and the rate of offending are assumed to be constant over age. In Chapter 4 we investigated serious, less serious, and trivial offenders. We showed how our simplified two category model adequately deals with particularly serious offences and approximately fits all serious offences. We also investigated the specialization and versatility of offenders and showed that the tendency is towards versatility rather than specialisation. We then extended our theory by the introduction of a category of offenders who commit mostly trivial offences and showed how a one category model could be used to explain the age profile of non-standard list convictions and make estimates of the recidivism probability \( p \) and conviction-rate parameter \( \lambda \) for trivial offences. These parameter values depended on the assumed size of the trivial offender category which may or may not include the high-risk/low-rate category of standard list offenders.

Introduction

The principal evidence in favour of our theory, and the models derived from it, is that we can reproduce the age–crime curve (the graph of the number of offenders convicted at any given age), both overall and for each conviction number first, second, third, and so on. This is despite the fact that our theory and models assume no causal relationship between the age of offenders and their criminal behaviour. The fall in the number of convictions at older ages is explained, at least until offenders become too infirm to offend, by
older offenders, on average, having been convicted more often, being more likely to have ‘retired’ from offending, with probability $(1 - p^n)$. The rise in the number of convictions during the early teenage years is explained by the progressive change from informal sanctions for antisocial/criminal behaviour, through police warnings and cautions to the ‘last resort’ use of formal prosecution and conviction in the magistrates’ or Crown courts.

More conventional theories (discussed in Chapter 1) either assume that individual age–crime curves are similar to the aggregate age–crime curve or that both onset and desistance are age-dependent. Certainly the former types of theory and many of the latter types also assume that the rate of offending $\lambda$ varies both between individuals and over time, $\lambda$ increasing to a peak in the teenage years and then decreasing as offenders get older. Our theory is inconsistent with these assumptions and in this chapter we set out to test whether the cohort data can be explained by these types of variable-$\lambda$ and/or age-dependent theories.

Generally, in our models, we have considered that the antisocial behaviour of young active offenders, over the period when convictions are increasing with age, has stayed constant. It is however possible that some offenders may desist from crime as a result of informal sanctions, reprimands, formal warnings, or cautions. As ever we must be careful to say that our statements refer to offending which could lead to conviction. There may well be age-dependent effects in the nature and rate of offending which do not show up in convictions. For example the nature of the crimes committed by

![Figure 5.1a Age–crime curve 1997 sentencing sample](image-url)
offenders may well change as they get older or progress further into a criminal career (LeBlanc and Frechette 1989; Piquero et al 2003; Tarling 1993). Despite what we believe to be compelling evidence for our theory, a critic might wonder if it is possible to fit the Offenders Index data with a more intuitive theory suggesting that offending behaviour, as measured by convictions, is causally dependent on age. We will now consider this possibility.

**Possible Types of Age Dependence**

A typical age–crime curve\(^1\) is shown in Figure 5.1a, in this instance for the sample of offenders convicted in 1997. The sample was drawn from the Offenders Index of all those convicted and sentenced during the first week of alternate months from February through to December 1997. The data plotted in Figure 5.1a has also been standardized (age-weighted) to a fixed number of individuals of each age in the community. This is very similar to the age–crime curve for the 1953 cohort shown in Figure 5.1b.

---

\(^1\) There are some features of the age–crime curves in Figure 5.1 which need further explanation. In Figure 5.1a there is a small secondary peak which is caused by the recording of unknown age as 25. In Figure 5.1b there is a step rise in crime during the teenage years at age 16 caused by the introduction of formal cautioning by the Metropolitan Police in 1968. Many offences at age 16 in 1969 led to unrecorded cautions rather than convictions, whereas offences at age 17 in 1970 (17 was the minimum age for the adult court) led to conviction. There is also an apparent earlier onset age in the 1953 cohort than in the 1997 sentencing sample, a point that we shall return to later.
One possible explanation for this similarity is that offenders are homogeneous in nature and that the aggregate curve directly indicates the probability of offending of each individual at each age (Gottfredson and Hirschi 1990). We will call this, and similar theories, ‘variable $\lambda$’ theories, since they imply that the individual frequency of offending varies with age. In this kind of theory it is age itself, acting as an average measure of psychological and social development (eg self-control) which determines the probability and rate of offending. The age–crime curve for an individual is assumed to be similar to the aggregate curve.

A second kind of explanation suggests that the rate of offending $\lambda$ over a given period is in fact constant from the ‘age of onset’, when offending starts, to the ‘age of desistance’ when offending ceases (Blumstein et al 1986). Each individual has a fixed career duration. By choosing suitable probability distributions of the ages of onset and desistance, when averaged over all offenders each with their individual onset and desistance ages, it should be possible to reproduce the aggregate age–crime curve. For example, Shinnar and Shinnar (1975), in their calculations of the incapacitative effects of custody, explicitly assumed that $\lambda$ was constant over age and that career length was exponentially distributed. Figure 5.2 shows a hypothesized criminal career for one individual. Each individual in the population of offenders would have their own similar career pattern but with a different $\lambda$, onset age and desistance age. Thus, the aggregate age–crime curve is very different from each individual age–crime curve. We will call these ‘fixed career length’ theories.

Obviously variable $\lambda$ and fixed career length theories fall at extreme ends of a whole range of theories, with varying probabilities of offending for a given individual at a particular age on the one
hand and varying onset and desistance distributions for the beginning and end of offending on the other. This may be further complicated by the existence of hybrid theories consisting of multiple developmental trajectories with differing parameters (Sampson and Laub 2005). However, if we can show that no theory or model with the features of either a variable $\lambda$ theory or fixed career length theory can fit the Offenders Index data, then we have ruled out all but the most contrived theories in which age is the main determinant of offending behaviour.

**Testing the Theories**

We begin by considering variable $\lambda$ theories. Here there exist one or more categories with varying rates of conviction at different ages. Each category in this sense has a particular relationship between age and the probability of conviction. To study this we can look at offenders who in their lifetime gather a large number of convictions, say more than five. Looking at the inter-offence times between successive convictions will show how their offending rate (measured in convictions) changes with age. In our models we have explained the behaviour of these prolific offenders by suggesting that they tend to be dominated by two high-risk categories, one with a fixed recidivism probability and a high offending rate and a second group with the same recidivism probability (for standard list offences) and a considerably lower rate of offending.

A pure ‘variable $\lambda$’ theory explains this as a group of crime-prone individuals each with their own propensity to offend which varies with age according to the age–crime curve. Gottfredson and Hirschi (1990) maintain the invariance of the age–crime curve and attribute the decline in adult crime with age to maturation processes and increasing self control slowing down the offending rate. The main point is that, if variable $\lambda$ theories are true, we should have a group of offenders over the age of 18 who, as they are getting older, would have a reducing rate of offending and, as a consequence, increasing inter-conviction times.

The conviction rate $\lambda$ (which we have defined as the conviction rate between onset and desistance) is usually measured in quantitative research as the number of convictions or arrests in some time period divided by the number of offenders. This measure is problematic in a number of ways: firstly, if the time period is short, say one year, then what is the correct divisor? Is it the total number of
Is Age the Primary Influence on Offending?

offenders in the group, or the number who are actually arrested or convicted in the time period? In the first case, desisters would be included in the denominator while in the second case, active offenders who happened not to be arrested or convicted in the time period would be excluded. It is also generally assumed that, for individuals, crime counts follow a Poisson distribution with mean $\lambda$, which implies that inter-conviction times are exponentially distributed with mean $1/\lambda$.

In Chapter 2 we saw that there is a graphical way of displaying the rate of conviction which overcomes the above measurement problems. We plot (with a logarithmic scale on the $y$ axis) the numbers surviving a specified time between consecutive convictions. If this plot is a straight line, the slope of the line gives a measure of the rate of offending. Indeed, it is the mean offending rate of the associated exponential distribution. As an example, a group of offenders about to be convicted for the seventh time will, on average, be at least a year older than those being convicted for the sixth time. If we therefore plot the inter-conviction survival time distribution for different offence serial numbers (for offenders aged over 18 on conviction) the change in the slope of the distributions will indicate any changes in the rate of offending with age.

What would we expect to see given a variable $\lambda$ theory? As offenders with higher conviction numbers are on average older, they should have progressively lower rates of offending. The slopes should therefore decrease as we go to higher conviction numbers. This is indicated in Figure 5.3. The graph shows the postulated survival time distributions of inter-offence times from the fifth to sixth, sixth to seventh conviction (etc), assuming a variable $\lambda$ theory. The graph is as usual on a logarithmic $y$ axis scale. The graph is based on a cohort size of 2,000 with a mean inter-conviction time of 1.1 years between the fifth and sixth convictions, increasing by 15 per cent for each subsequent conviction number.

What actually happens, in the Offenders Index 1953 cohort, for offenders over the age of 18 at increasing stages of the criminal career, is shown in Figure 5.4. The lines plotted in Figure 5.4 are not ‘best fit’ curves but are the actual survival time data from one conviction to the next for the specified conviction serial numbers. The uppermost curve is the combined inter-conviction survival times from first to second, second to third, third to fourth, fourth to fifth, and fifth to sixth appearances. The lower curves are for the inter-conviction survival times for the pairs of conviction serial numbers.
Testing the Theories 113

Figure 5.3 Hypothetical survival time plots for increasing conviction number

(AppNo) specified in the legend, with the steeper curve combining all inter-conviction survival times to conviction serial numbers greater than the twelfth.

For conviction numbers (AppNo) between the sixth and the twelfth the survival time curves are essentially parallel, indicating constant survival time distributions differing only in the size of the subsets. The survival curve for convictions up to and including the sixth again appears to be parallel. For convictions over the twelfth

Figure 5.4 Survival time curves for the 1953 cohort (Age at conviction >18)
the curve is steeper than the other curves. This suggests that for higher conviction numbers, and therefore older offenders, the inter-offence times apparently get shorter and offending speeds up. However, under our theory we would expect this apparent speeding up because of the cut-off point of the cohort data at age 46. Fitting the double exponential survival functions derived in Chapter 2 confirms that the parameters are essentially constant as the conviction serial numbers increase.

Figure 5.5 plots the parameter values for the ‘best fit’ double exponential survival functions to the curves in Figure 5.4. As our theory predicts, the mean number of convictions per year for the high and low-rate parts of the curves, parameters $\lambda_1$ and $\lambda_2$ respectively, remain essentially constant as the conviction serial numbers increase. The anticipated increase in the rate for serial numbers above 12, due to censoring, is also evident. There is no evidence that the slopes get shallower, ie that offending is slowing down.

For a variable $\lambda$ theory to adequately explain the age–crime curves the rate parameters would need to be reduced by about 15–20 per cent between each offence. The solid horizontal lines in Figure 5.5 are the parameter values for the complete 1953 cohort.

---

2 The parameters $\lambda_1$ and $\lambda_2$ are equivalent to the $\lambda$s used in other quantitative studies but are estimated directly from the slopes of the inter-conviction survival time curves and not by counting convictions in given time periods.
The proportion (high-rate) parameter ‘b’ for conviction serial numbers under seven is significantly less than the complete cohort value, indicating the preponderance of low-risk/low-rate offenders in this group. For conviction serial numbers over six, b is higher than the complete cohort value.

The above analysis demonstrates that at any stage in the criminal career the distribution of time to next conviction is invariant and not dependent on prior criminal history. As a further illustration, if we look at individuals who in their lifetime accrue eight or more convictions, our theory would predict that early convictions, those with serial numbers less than eight, would have rate parameters consistent with both those for serial numbers greater than eight and those of the cohort as a whole.

Table 5.1 compares the ‘best fit’ parameter values for inter-conviction survival times for offenders with at least eight convictions in their criminal career, separately for later convictions, with serial numbers greater than seven and for earlier convictions with serial numbers less than eight. From Table 5.1 we can see that the proportion of high-rate offenders is higher for later convictions, in line with our theory’s prediction that low-rate offenders will be less likely to reach the higher conviction numbers. Although both rate parameters are lower for the later convictions, suggesting a slight slowing down in offending, the difference is very small and cannot account for the observed age–crime curve. We can thus rule out variable λ theories for all those offenders who commit standard list offences.

This leaves us with fixed career duration theories, in which the offenders offend at a constant rate but their onset and desistance ages are given by some distributions. Shinnar and Shinnar (1975) used a negative exponential distribution for residual career length in their incapacitation calculations. We observe in the Appendix

<table>
<thead>
<tr>
<th>Conviction number</th>
<th>Proportion high-rates: b</th>
<th>High-rate parameters: λh</th>
<th>Low-rate parameters: λl</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Late) &gt;7</td>
<td>0.70</td>
<td>0.95</td>
<td>0.30</td>
</tr>
<tr>
<td>(Early) &lt;8</td>
<td>0.63</td>
<td>0.98</td>
<td>0.38</td>
</tr>
</tbody>
</table>
that the average residual career length of both the high and low categories, who commit offences serious enough to be sent to prison, is around 5.3 years (the high category being a little shorter, the low category a little longer). The Offenders Index data that we have considered so far could therefore be approximately modelled by a fixed career length theory with two categories of offenders with constant (but different) rates of conviction and a (negative exponential) distribution of career length with a mean of 5.3 years.

This model could theoretically be tested in an experiment in which we have two groups of serious offenders convicted over the same period. We take one group out of circulation for a period of say about seven months; the other group is immediately put back into society so that they can offend again. The recidivism of the two groups is then compared over a fixed period at liberty (say two years) having corrected for the seriousness of offences originally committed by the two groups. Because the group who could not offend for seven months will be on average seven months older, a proportion \(1 - e^{-7/(5.3 \times 12)}\) = 10.4 per cent would have reached the end of their offending careers and we should therefore see a reduction in the proportion reconvicted of 10.4 per cent. But we would only expect about half of the reconvictions to occur in the two year follow-up period, resulting in a 5 per cent reduction in the two-year reconviction rate for custody compared with non-custodial disposals. This experiment is similar to experiments comparing the effects of custodial and non-custodial sentences (see eg Killias et al 2010).

We now look at one such study based on the Offenders Index. This is described by Kershaw (1999). The two-year reconviction rates of offenders discharged from prison over the period 1987 until 1996 were compared with those receiving community penalties. After correcting for seriousness of offence and so called ‘pseudo-reconvictions’,\(^3\) no discernable difference (−0.7 per cent, 1995, and −1.3 per cent, first quarter of 1996) can be seen in the reconviction rates; certainly less than the −5 per cent predicted for the average of seven months spent in prison by those receiving custodial sentences. The correction for pseudo-reconvictions has raised issues about the validity of drawing conclusions from Kershaw’s study about the additional individual deterrent effect of a prison sentence. However, with no correction, the two year

---

\(^3\) A ‘pseudo-reconviction’ is where an offender is subsequently (re-)convicted for an offence committed prior to the date of the current sentence.
reconviction rate for community penalties is the same as that for imprisonment, which provides no support at all for a fixed career duration theory.

Ideally we would like to look at total career recidivism following custodial and non-custodial sentences, but recidivism rates following specific disposals may not be reliably calculable from a cohort sample. This is because the nature of disposals may change over the time span of the cohort, particularly for non-custodial sentences. It could also be argued that, as the same offenders reappear many times in the sample of conviction occasions, the results are in some way biased. However, despite these reservations we now assume that the effectiveness of custody and community penalties did not change over the observation period. We can also control for seriousness and type of offender by looking at the subset of offenders who have at least one custodial sentence and more than four convictions.

In the 1953 cohort, 1,141 offenders fall into this subset, accruing a total of 7,016 convictions; 2,561 of these resulted in custody and 2,172 of these were followed by a further conviction. The overall reconviction rate for custodial sentences was therefore 84.8 per cent. In the same subset of offenders, of the 4,455 convictions not resulting in custody 3,703 were followed by further convictions and the overall reconviction rate for non-custodial sentences was 83.1 per cent. Both of these reconviction rates are consistent with the high-risk reconviction probability of 0.84 estimated for the whole 1953+ cohort (see Table 2.1). Again we see no evidence of the reduced recidivism following custody that is predicted by a fixed career length theory.

It could be argued that both cohort analysis (see above) and short follow-up times are unreliable in detecting the expected reduction in recidivism due to career termination whilst in custody. We therefore need to use a longer follow-up time cross-sectional analysis to directly estimate recidivism for different disposals on a more consistent basis. The follow-up period usually used in reconviction studies is two years; this period is long enough to allow a substantial proportion to be reconvicted but short enough not to appear to be out-of-date in policy terms. However, we have shown from our earlier analysis that reconviction times, for a significant proportion of cohort members, are very much greater than two years.

We can replicate Kershaw’s (1999) analysis using the 1997 sentencing sample, drawn from the Offenders Index, with a much
longer follow-up period. Table 5.2 shows the numbers and proportions reconvicted during a follow-up period (to 31 December 2001) of those convicted of standard list offences during the first week of alternate months from February through to December 1997 (six weeks in all). Where individuals were convicted in more than one of the sample weeks the earliest conviction has been taken as the target conviction for the analysis. No corrections have been made for pseudo-reconvictions or time served in custody, either on remand before conviction or under sentence during the follow-up period. The follow-up period varies between four years 11 months for the earliest convictions to four years and three weeks for the latest convictions, but for the purpose of the analysis only, reconviction times less than four years three weeks have been counted as reconvictions.

It can be seen from Table 5.2 that those with custodial sentences had the highest reconviction rate at 64.4 per cent. Sentences involving supervision in the community had a similar rate at 62.5 per cent but fines and other disposals had significantly lower rates at 43.4 per cent and 51.5 per cent respectively. However, it seems likely that different disposals are given to different kinds of offender, and it is important to control for these differences. As a first step we can assume that the magistrates and judges take account of at least some of the characteristics of the offenders and aspects of their criminal careers in making their sentencing decisions. Supervision and custody are often regarded as alternatives for the lower end of more serious offending and should therefore be directly comparable. However, the difference in recidivism probabilities between them (which is significant at $p = 0.01$; $\text{Chi}^2 = 9.15$ on $1df$) is in the wrong direction to support a fixed career length theory.

Table 5.2 Proportion of offenders reconvicted during a 4 year 3 week follow-up period for the 1997 sentencing sample

<table>
<thead>
<tr>
<th></th>
<th>Not reconvicted</th>
<th>Reconvicted</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine</td>
<td>13143</td>
<td>56.6%</td>
<td>10077</td>
</tr>
<tr>
<td>Supervision</td>
<td>5534</td>
<td>37.5%</td>
<td>9224</td>
</tr>
<tr>
<td>Custody</td>
<td>3629</td>
<td>35.6%</td>
<td>6559</td>
</tr>
<tr>
<td>Other</td>
<td>4882</td>
<td>48.5%</td>
<td>5189</td>
</tr>
<tr>
<td>Total</td>
<td>27188</td>
<td>46.7%</td>
<td>31049</td>
</tr>
</tbody>
</table>
If we exclude first convictions from the comparison the difference in recidivism probabilities between custody and supervision diminishes to less than 1.2 per cent (which is not significant at \( p = 0.05 \), \( \chi^2 = 2.94 \) on 1 df). Excluding more of the early convictions further diminishes the difference in reconviction rates between supervision and custody, but the rates themselves increase with each successive conviction omitted from the comparison. With conviction numbers greater than six the proportions reconvicted after supervision or custody are 82.3 per cent and 82.9 per cent respectively (\( \chi^2 = 0.51 \) on 1 df, not significant). If adjustments were to be made for pseudo-reconvictions and/or time spent in custody during the follow-up period, these would have the effect of widening the gap by reducing recidivism for supervision and increasing it for custody. As with Kershaw’s analysis, this later sample with a longer follow-up period provides no evidence of age related desistance or lower reconviction rates following custody and no support for fixed career duration theories.

As a by-product of his calculation of the incapacitative effect of a prison sentence, Tarling (1993, p 145) showed that the assumption that recidivism probabilities of offenders in the Offenders Index are generated by a fixed career length theory, implies that the average residual career length would be about two to three years. We have seen that a direct analysis of the Offenders Index shows that the average is nearer five years for high-risk/high-rate offenders and 10 to 15 years for high-risk/low-rate offenders. Again there is no support for fixed career length theories in the Offenders Index data. The lack of any unequivocal evidence for such a reduction in recidivism (McGuire 1995) suggests that we can rule out fixed career length theories.

In general, then, we can rule out any theory (variable \( \lambda \), fixed career length or any mixture) in which offending behaviour is dominated by the effects of age, at least for those offenders who commit standard list offences, the high and the low-offending rate categories. This is not to say that offending does not depend on age at all, only that: offenders in these categories do not desist from crime only or mainly because they are getting older.

**Conclusion**

It is clear from the many studies of age and crime that there is an aggregate relationship between age and criminal offending.
Our theory, derived in Chapters 2 and 3, explains this relationship as the result of a process of capture, conviction, and sentencing of individuals. Homogeneous categories of individuals have constant probabilities of offending in given time periods and have constant probabilities of recidivism after each conviction. The only part of our theory which is age-dependent is the apparent initial rise in crime during the early teenage years, when society’s response to criminal behaviour changes with age. The largely informal response at the age of 10 becomes progressively more severe with an increasing proportion of offenders subjected to full criminal proceedings as they approach and enter adulthood.

Our criminal career theory fits the Offenders Index data without assuming that adult offending (convictions) depends on age. The decline in crime with age in the adult population is entirely explainable by the constant (within each group) proportion of offenders desisting after each conviction. The apparent reduction in the frequency of offending of older offenders is due to the increasing proportion of low-rate offenders in the active population as age increases. There is no support in the Offenders Index data for any change in the (within Category) rate of offending with age.

In testing the fixed career duration type of theory, in which each individual has his or her own fixed career length which when aggregated over all individuals produces the age–crime curve, no reduction in recidivism was observed after custodial sentences. This result is consistent with our proposition that desistance occurs at the time of conviction and that on release prisoners, who have not desisted, simply rejoin the active offender population. This also implies that there is no major reduction in crime caused by keeping people in prison, given the current average time served. The only way to create a reduction in crime by incapacitation is to constantly increase the prison population and/or keep offenders in prison for much longer periods. We will return to this topic in our discussion of the significance of our theory for criminal justice policy in Chapter 8 and in the Appendix.

One important point should be made here. Age-based theories tend to assume, like Shinnar and Shinnar (1975) and Gottfredson and Hirschi (1990), that offenders grow out of crime and that the effects of the criminal justice system on offending are minimal, except through incapacitation. In our theory custody in itself ‘does not work’ (either by individual deterrence or incapacitation, except for those who are removed for very long periods from society).
There is no support for age-based desistance in our analysis as we have shown that the expected residual career length for a recidivist is exactly the same when released from prison as it is when leaving court with a non-custodial sentence.

This chapter has explored alternative age-based theories and has shown that theories that assume explicit age dependence cannot fit the data. However, our theory does suggest that the operation of the Criminal Justice System, in repeatedly convicting offenders, is the most important factor in reducing the number of active offenders and is essential in the control of crime.