General conclusion and discussion
In criminology, little is known about non-arrested offenders because data sources that are traditionally used do not offer many opportunities to study this group of offenders and their offending behaviour. In this thesis an alternative data source is used to study the offending behaviour of non-arrested offenders. Data from the Dutch DNA database offer the opportunity to study some aspects of the offending behaviour of non-arrested offenders, compare this to the offending behaviour of arrested offenders, and with that to study whether characteristics of offending behaviour have an influence on the probability that an offender will be arrested.

The purpose of this thesis is to study non-arrested offenders and characteristics of their offending behaviour and compare these to the characteristics of the offending behaviour of arrested offenders. The main question of this thesis therefore is: Which characteristics of offending behaviour influence the probability that an offender will be arrested? This thesis focuses on two main types of characteristics: criminal career characteristics and spatial characteristics. The main question of this thesis is therefore divided into two sub questions: (1) Do criminal career characteristics influence the probability that an offender will be arrested?; and (2) Do spatial characteristics of offending behaviour influence the probability that an offender will be arrested? The comparison of the offending behaviour of non-arrested offenders with that of arrested offenders enables us to draw conclusions on whether characteristics of this offending behaviour have an influence on the probability that an offender will be arrested. Consequently, conclusions can also be drawn on how representative official arrest data are, on which many criminological studies are based: if certain characteristics have an influence on the probability that an offender ends up in arrest data, these data are not formed by chance and are thus selective when it comes to these characteristics. Offenders with these characteristics are over- or underrepresented in arrest data.

The main question of this thesis is answered by summarizing the results of the previous chapters in section 5.1. This section will also briefly discuss what this answer implies for the representativeness of arrest data (a more elaborate discussion of this topic can be found in section 5.5). In section 5.2 the limitations of this thesis will be discussed, specifically focused on the selectivity of DNA data, and in section 5.3 more general limitations of using DNA traces for criminological research will be discussed. Section 5.4 describes future research possibilities and in the last section (5.5) policy implications of this thesis are discussed, including the implications for studies that use arrest data to study either criminal careers or spatial offending behaviour.
5.1 Do characteristics of offending behaviour influence the probability that an offender will be arrested?

The main question of this thesis can be answered by answering the two sub questions. This is done by summarizing the results of chapters 2, 3 and 4. Chapter 2 deals with the influence of criminal career characteristics on the probability that an offender will be arrested, and thus answers the first sub question. The second sub question in answered in chapters 3 and 4. Chapter 3 studies whether the spatial dispersion of crime locations of an offender on a meso-level influences the probability of arrest, by analysing whether the number of police regions in which an offender commits his crimes influences this probability. Chapter 4 studies whether micro-level spatial dispersion of crime locations of an offender (or the criminal range of an offender), measured as the average distance between his crime locations, influences the probability that he will be arrested.

The theoretical part of chapters 2 and 3 both describe that the clearance rate (the rate of crimes reported to the police for which an offender is arrested) in many Western countries is low. Low clearance rates are often seen as having a negative effect on the legitimacy of the criminal justice system (i.e. Litwin & Xu, 2007; Roberts, 2007). Also, both specific and general deterrence depend on the certainty and swiftness of punishment (i.e. Blumstein, Cohen, & Nagin, 1978; Nagin, 1998). Unsolved crimes may make offenders think they can escape justice and continue to commit crimes (Paré, Felson, & Ouimet, 2007). It is therefore important to not only know what influence the likelihood that individual crimes are solved, but also what characterizes offenders who continue to escape arrest.

Chapter 2 gives insight into this by studying whether criminal career characteristics have an influence on the probability of arrest, and hereby answers the first question of this thesis. Three characteristics of the criminal career are studied in this chapter: (1) the length of the criminal career (measured as the number of committed crimes), (2), the seriousness of the committed crimes and (3) offense specialization. The hypothesis about the number of crimes is that as an offender commits more crimes, the probability that he will be arrested increases. Every crime in itself carries a certain probability of being cleared, and consequently an offender who commits multiple crimes has a greater likelihood of being arrested than an offender who has committed one crime, due the combined probability of arrest of the separate offenses. The second hypothesis described in chapter 2 is that as the seriousness of the crimes committed by an offender increases, so does his probability of arrest. The police and the criminal justice system will spend more time and effort in solving serious crimes like homicide and other violent offenses, than in solving less serious crimes like burglary. The last hypothesis tested in this chapter is that specialized offenders have a smaller probability of being arrested than generalists. More experience in
committing a certain type of crime is likely to lead to more expertise, which may lead to a smaller probability of arrest (i.e. Clare, 2010; Topalli, 2005).

Chapters 3 and 4 answer the second sub question of this thesis, as they both focus on the influence of spatial characteristics of offending behaviour on the probability of arrest. Chapter 3 focuses on a meso-level spatial characteristic: the number of police regions in which an offender commits crimes (see figure 5.2a for the 25 Dutch police regions). Egger (1984, 1990) and Rossmo (2000) describe that there is a lack of sharing and coordinating of investigative information, and issues of cooperation and competition between different police regions (whether these regions are actually jurisdictions or just separate geographical areas). Therefore the main hypothesis in this chapter is that if an offender commits his crimes in multiple police regions, the probability of him getting arrested is smaller than if he commits crimes in one region. In this chapter also the distance between regions is studied, as well as the spatio-temporal ordering of crimes (how many times an offender changes regions). For both measures the hypothesis is the same: when the distance or the number of times an offender changes region increases, the probability of arrest decreases.

Chapter 4 focuses on a micro-level spatial characteristic of offending behaviour: the criminal range of an offender. This chapter has a somewhat different theoretical point of view compared to chapters 2 and 3. This chapter describes that many different authors who study the spatial behaviour of offenders have mentioned that their results and results of similar studies may be biased because only arrested offenders are studied (see for example Canter, Coffey, Huntley, & Missen, 2000; Eck & Weisburd, 1995; Elffers, 2004; Goodwill & Alison, 2005; Johnson, Summers, & Pease, 2009). Therefore this chapter compares the spatial dispersion of crime locations (the criminal range) of arrested offenders to the spatial dispersion of crime locations of non-arrested offenders. Spatial dispersion, or the criminal range, is measured as the average distance between the crime locations in a series. Some of the previously mentioned authors also commented on the fact that offenders may be arrested because of their spatial behaviour: offenders with a limited range of operation (small average distance between crimes) may have a greater probability of being arrested. This may causes selectivity in arrest data and bias in studies based on this data (Eck & Weisburd, 1995; Johnson et al., 2009). This chapter therefore also analysed whether the spatial dispersion of crime locations has an influence on the probability of arrest.

In all chapters survival analysis, and more specifically Cox proportional hazards model (Cox, 1972) was used to analyse whether the offending behaviour characteristics have an influence on the probability of arrest. Survival analysis is used in these chapters for two reasons: (1) survival analysis studies how long it takes for an event of interest to take place, given that the individual is still at risk for experiencing the event. It does therefore not only take into account whether an offender is arrested or not, it also considers how
long it took for the offender to become arrested. And (2) survival analysis efficiently
utilises censored durations, which means that it recognises the possibility that the arrest of
the offender can still happen after the end of the study period.

The outcome measure of a survival analysis is a hazard ratio. The hazard ratio is an
indicator of the effect of the independent variable on the hazard of the event of interest. The
hazard ratio can be interpreted as the change in the probability of experiencing the
event of interest that is the result of a one-unit change in the independent variable (Cleves,
Gould, Gutierrez, & Marchenko 2008). In this thesis, the event of interest is the arrest of
the offender and the independent variables are the different characteristics of the offending
behaviour. The hazard ratio can thus be interpreted as change in the probability of arrest as
a result of a one-unit change in these variables.

In Table 5.1 a, b, and c the results of the analyses of the different chapters are shown.
In all these analyses the dependent variable was the time until arrest, and the independent
variables are the different offending behaviour characteristics.

<table>
<thead>
<tr>
<th>Table 5.1 a</th>
<th>Results analyses chapter 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td></td>
<td>HR</td>
</tr>
<tr>
<td><strong>Criminal career characteristics</strong></td>
<td></td>
</tr>
<tr>
<td># Committed crimes</td>
<td>1.19***</td>
</tr>
<tr>
<td>Seriousness</td>
<td>0.99</td>
</tr>
<tr>
<td>Specialization (Diversity Index)</td>
<td>2.84***</td>
</tr>
<tr>
<td>Specialization in</td>
<td></td>
</tr>
<tr>
<td>Violent offenses</td>
<td>0.36***</td>
</tr>
<tr>
<td>Sex offenses</td>
<td>0.50</td>
</tr>
<tr>
<td>Burglary</td>
<td>0.86***</td>
</tr>
<tr>
<td>Theft</td>
<td>0.43***</td>
</tr>
<tr>
<td>Theft of/from car</td>
<td>0.83*</td>
</tr>
</tbody>
</table>
Two models are analysed in chapter 2 (see Table 5.1 a). In the first model specialization is measured using the diversity index, which is a measure that does not take into account the particular crime type an offender is specialized in. The second model analyses crime-specific specialization measures. In both models the number of crimes committed and the seriousness of the committed crimes are analysed, and in both models the results for these two variables are the same: the number of crimes has a significant hazard ratio above 1, indicating that as the number of crimes an offender commits increases, so does his probability of arrest; the seriousness of the committed crimes does not have a significant influence on the probability of arrest.

The diversity index as a measure of specialization shows that a completely versatile offender (score of 1 on the diversity index) is 2.84 times more probable to get arrested than a totally specialized offender (score of 0 on the diversity index). When specialization in different types of crimes is taken into account in the second model analysed, offenders who specialize in violent crimes, burglaries, thefts and car thefts all have a smaller

---

**Table 5.1 b  Results analyses chapter 3**

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR</td>
<td>SE</td>
<td>HR</td>
<td>SE</td>
</tr>
<tr>
<td># Police regions</td>
<td>0.91**</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Distance between regions</td>
<td>0.93**</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td># Changes between regions</td>
<td>0.87***</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td># Police regions</td>
<td>0.88**</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Distance between regions</td>
<td>1.01</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>

**Spatial characteristics:**

**Police regions**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td># Police regions</td>
<td>0.91**</td>
<td>0.03</td>
<td>0.88**</td>
<td>0.03</td>
</tr>
<tr>
<td>Distance between regions</td>
<td>0.93**</td>
<td>0.02</td>
<td>0.79**</td>
<td>0.01</td>
</tr>
<tr>
<td># Changes between regions</td>
<td>0.87***</td>
<td>0.02</td>
<td>0.72***</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Table 5.1 c  Results analysis chapter 4**

<table>
<thead>
<tr>
<th>Spatial characteristics:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal range</td>
</tr>
<tr>
<td>Average distance between crimes</td>
</tr>
</tbody>
</table>

HR: Hazard ratio; SE: Standard Error. *p<.05; **p<0.01; ***p<0.001
probability of being arrested than versatile offenders. The answer to the first sub question of this thesis is thus that the length of the criminal career, measured as the number of committed crimes, as well as offense specialization influences the probability than an offender will be arrested. This also means that offenders who are arrested, and are therefore present in official arrest data, are different in a number of respects from offenders who escape arrest and who are not present in arrest data. Offenders present in official arrest data are more often generalists and have committed more crimes than offenders who are not present in official arrest data (section 5.5 will discuss this more elaborately).

In chapter 3 three measures of meso-level spatial dispersion of crime locations of an offender are analysed. These are the number of police regions in which an offender commits his crimes, the distance between these police regions, and number of times an offender changes between police regions to commit crimes (the spatio-temporal ordering of the committed crimes). These three measures were first analysed in three separate models, see Table 5.1 b. The results of these models show that all three measures have a significant hazard ratio below 1, indicating that as each of these measures increases, the probability of arrest decreases. After analysing these measures separately, the question remained whether the distance between police regions in which an offender commits crimes, or just the number of police regions in which an offender commits crimes causes this decreased probability of arrest. This is analysed in the fourth model of chapter 3, by analysing both measures (controlling for the number of committed crimes) in one model. Model 4 in Table 5.1 b shows that it is actually the number of regions in which an offender commits his crimes, and not the distance between these regions that causes the decreased probability of arrest; the hazard ratio of the distance is no longer significant in this model.

Chapter 4 analyses the micro-level spatial dispersion of crime locations of an offender. In this chapter it is studied whether the average distance between crimes locations influences the probability of arrest. The results show that this average distance is unrelated to the probability of arrest (see Table 5.1 c).

Therefore the answer to the second sub question is that the number of police regions in which an offender commits his crimes does have an influence on the probability of arrest, but that the average distance between crimes committed by one offender does not. This means that offenders who are studied when arrest data are used are more often offenders who commit their crimes in one police region. Also, when arrest data of one police region are used, one has to keep in mind that offenders who operate less locally and tend to cross region boundaries, might not be present in the same number in these data as the offenders who only committed their crimes in that one single region. However, because the average distance between crime locations does not influence the probability of arrest, arrested offenders are probably representative when it comes to the average distance
between crime locations. Concerns raised by many researchers (see before) that arrested offenders are likely not representative in terms of their spatial behaviour, might not be justified when it comes to the average distance that offenders travel between crime locations.

The answer to the main question of this thesis “Which characteristics of offending behaviour influence the probability that an offender will be arrested?” is therefore that the number of crimes an offenders commits, offense specialization, and the number of regions have an influence on the probability that an offender will be arrested. The more crimes an offender commits, the greater the probability that he will be arrested, specialized offenders have a smaller probability of being arrested than generalists, and as the number of police regions in which an offender commits his crimes increases his probability of being arrested decreases. Again, the implications of these results for the use of arrest data for criminological research will be discussed in section 5.5.

Another conclusion from this thesis is that DNA traces do have an added value for criminological research. Tilley and Townsley (2009) stated that “by using forensic data in an analytical manner, it should be possible to better understand the size, nature, structure and dynamics of the offender population” (p. 15). This thesis is a good example of that.

5.2 Selectivity of DNA data

This thesis has clearly shown that DNA traces have an added value for criminological research. Currently DNA traces are rarely used for this type of research, but they do have potential for future research. However, little is known about the selectivity of DNA data itself. DNA data are police data; they are the product of decisions at various stages of the investigative process and therefore DNA data potentially suffer from the same problem as arrest data: selectivity. Like traditional types of police data they reflect the priorities and technical possibilities of the process of criminal investigation. According to Tilley and Townsley (2009), the selectivity of DNA data is different from both traditional types of police data (for example arrest data or records from prisons) and from self-report surveys, although they do potentially suffer from some form of bias. Compared to arrest data, DNA data are collected earlier in the investigative process and therefore suffer less attrition (Tilley & Townsley, 2009). It is important to have some insight into the nature and extent of the representativeness of DNA data, because it tells something about the generalizability of the results of studies that use DNA data, and thus also about the generalizability of the results of this thesis. This section will discuss studies that can give some insight into the selectivity of DNA data. A number of studies have been conducted on the collection of DNA traces at crime scenes, although these studies were not
specifically focused on the selectivity of data. They are mostly conducted from a forensic point-of-view.

Before a DNA trace is loaded onto the DNA database, the first police officer on the crime scene has to include the forensic department in the investigation, a crime scene examiner (CSE) from the forensic department has to visit the crime scene, and a DNA trace needs to be secured at the crime scene by a CSE. Previous studies focused on two steps of this process: (1) the visiting rate: the percentage of crime scenes visited by a CSE, and (2) the retrieval rate: the percentage of crime scenes yielding a DNA trace. This section of the discussion will focus on the representativeness of DNA data in terms of different crime types, and in terms of different police regions: do different crime types and different police regions have different visiting and/or retrieval rates? Previous studies mostly focused on crime types and police regions, and both these themes are also highly relevant for the previous chapters of this thesis.

Type of crime
Tilley and Townsley (2009) discuss that a clear distinction can be made between the collection of DNA traces at crime scenes of serious but relatively rare crimes and the collection of DNA traces at crime scenes of high volume crimes, such as burglary. The former is an intensive process, while the latter is done routinely. This is the result of the amount of effort that law enforcement is willing and able to put into solving a certain type of crime. Compared to the number of crime scenes from serious crimes like murder, the number of high volume crime scenes is much greater. There are not enough resources available to execute the speculative and thorough search for DNA traces that would be done on the crime scene of a major crime (for example on a homicide or rape crime scene) at every crime scene of a high volume crime (Tilley & Townsley, 2009). As a result, the search for traces at high volume crime scenes is less intensive and extensive and therefore the likelihood that DNA traces will be secured at crime scenes of high volume crimes is smaller. Peterson, Sommers, Baskin, and Johnson (2010) find that the collection of forensic evidence (including DNA traces) is very extensive at homicide crime scenes and to a lesser extent rape scenes, and that it is more limited for assault, burglary and robbery offenses. They show that in 0.5 per cent of assault cases biological evidence (including DNA) was submitted to the laboratory, for burglary cases this was 0.3 per cent, for homicide cases 32.3 per cent, for rape cases 22.6 per cent, and for robbery cases 0.7 per cent (Peterson et al., 2010). Burrows et al. (2005) find that inter-crime variations exist in retrieval rates: about 5 per cent of burglaries yield a DNA trace, compared to 6 per cent for theft of a motor vehicle and 1 per cent for theft from a motor vehicle.

Not many empirical studies on the collection of DNA traces at Dutch crime scenes exists, this section is therefore based on two studies, namely Jacobs and Bruinsma (2008), and
Goedvolk, van de Grift, and Huitink (2010). In the Netherlands a police officer from the uniformed police will arrive first at the crime scene after a crime is reported to the police. When the police officers that arrive first at the crime scene suspect that the incident may be a serious crime like murder, the forensic department will be called. If it concerns a high volume crime, like car theft, the police officers first on the scene will make the decision whether or not to involve the forensic department based on whether or not they think it is useful and necessary for a CSE to visit the crime scene (Goedvolk, van de Grift, & Huitink, 2010).

After the forensic department is informed by the uniformed police about a committed crime, they will make a decision whether or not to visit the crime scene. This decision will be made based on information the forensic department gets from the police officers on the scene about the crime scene and the circumstances. Crime scenes of serious crimes like murder will always be visited by a team of CSE’s. The policy in Dutch police regions is to let a CSE visit every burglary crime scene, unless there are strong reasons not to (for example because the house is already cleaned after the burglary). According to Goedvolk et al. (2010) most forensic departments will not visit a high volume crime scene if there are no clearly visible traces (Goedvolk et al., 2010). Goedvolk et al. (2010) show that in four Dutch police regions the visiting rates of residential burglaries lies around the 40 per cent, while the visiting rate of commercial burglaries lies around 20 per cent, so clearly not every burglary crime scene is visited by a CSE. Numbers on thefts of or from cars are often not registered separately, but these types of cases do not have a high priority in many regions, so the visiting rates of these types of crimes scenes are probably low (Goedvolk et al., 2010). Retrieval rates were not studied by Goedvolk et al., however Jacobs and Bruinsma (2008) find that 1.3 per cent of violent crimes yielded a DNA trace, as did 6.2 per cent of sex offenses, and 1.1 per cent of capital offenses.

Differences between police regions

As described, the first police officer(s) that attends the crime scene decide(s) whether or not to include the forensic department in the investigation. According to Peterson et al. (2010) different US police agencies can have different policies with respect to calling a CSE to the crime scene. Burrows et al. (2005) studied eight Basic Command Units (BCUs) in England and Wales, to examine general levels of attrition in volume crime cases. They find that differences exist between the visiting rates of the eight BCUs. For domestic burglaries the percentage of crime scenes visited by a CSE varies between 63.1 and 100 per cent, for non-domestic burglaries between 34.4 and 73.7 per cent, and when all volume crimes are considered together the visiting rate varies between 31 and 55.1 per cent (Burrows et al., 2005). Bradbury and Feist (2005) in their meta-analysis of the literature on the use of forensic science in volume crime investigations, state there are geographical variations between the number of crime scenes from which forensic traces are secured.
Goedvolk et al. (2010) also shows that differences exists between different Dutch police regions in visiting rates of crime scenes of both residential and commercial burglaries. They studied four Dutch police regions and find that the 2009 visiting rate for commercial burglaries varies between 31 and 54 per cent, and the visiting rate for commercial burglaries varies between 18 and 28 per cent (Goedvolk et al., 2010). This study also shows that there is a difference between Dutch police regions in the percentage of high volume crime scenes that is called in to the forensic department by the uniformed police. Some regions report around 50 per cent of crimes scenes to the forensic department, for other regions this is 90 per cent (Goedvolk et al., 2010). Jacobs and Bruinsma find that for four Dutch police regions the retrieval rates for all crime types together vary between 1 and 2.9 per cent. So although differences between Dutch police regions in both visiting and retrieval rates do exist, these differences seem not large.

Three important points regarding the limitations of this thesis need to be made based on this short literature review.

First, the overall retrieval rate is low: DNA traces are secured at the minority of crime scenes. This means that the crimes present in DNA data are only a small subsample of all crimes that are committed, which is not necessarily problematic if this subsample is representative for the entire sample.

Second, with regard to the different crime types, it is clear that burglaries are over-represented in DNA data compared to other crime types, because the policy in Dutch police regions is that a CSE needs to visit every burglary crime scene. Consequently more burglaries will be present in the data compared to other crime types, simply because of the higher visiting-rates. This could have consequences for the results of chapter 2, the chapter that analysed criminal career characteristics. The data used for chapter 2 consist for 68 per cent of burglaries, the percentage of registered burglaries in the Netherlands in 2009 (this year was chosen because the DNA data used for this thesis ends with the year 2009) is 6 per cent. This means that there is a large overrepresentation of burglaries in the DNA data compared to registered police data. This could mean that the burglary specialization measure used in chapter 2 is actually an artefact of the overrepresentation of burglaries. Consequently, the results of chapter 2 may only be valid for burglary specialization, but may not be generalizable to other crime specific specialization measures. However, the other specialization variables (both the diversity index and the other crime type specific specialization measures) show the same result as the burglary specialization variable in the survival analysis. Therefore we are inclined to believe that the results have not been affected too much.

The third important point is that there probably is a difference between police regions in the percentage of crime scenes visited, retrieval rates, and retrieval rates per crime type. The studies by Jacobs and Bruinsma (2008), Goedvolk, van de Grift, and Huitink (2010),
and Peterson, Sommers, Baskin, and Johnson (2010) do indicate that differences may exist. This could be problematic for the results of chapter 3, the chapter that studied spatial dispersion of crimes across police regions. The distribution of crimes across the 25 Dutch police regions in the DNA data used for chapter 3, shows that the minimum percentage of crimes in one region is 1.4 per cent of all crimes in the data, and the maximum is 10.2 per cent, so a difference does exist between regions in the number of committed crimes were a DNA trace was secured, although this difference is not too large. The percentages of different crime types across the different regions in the DNA data are very similar to the overall crime distribution across different regions. Whether the differences between regions stem from different percentages of crime scenes visited, different retrieval rates, or from other mechanisms is unknown. However, the distribution of registered crimes in 2009 (again, this year was chosen because the DNA data used for this thesis ends with the year 2009) across the different police regions (CBS Statline) shows a very similar distribution to the distribution in the DNA data. Figure 5.1 shows the comparison of the crime percentages of the 25 Dutch police regions, according to the registered data (x-axis) and the DNA data (y-axis). As can be seen, the distribution of crimes across the 25 police regions (represented by the black diamonds) is very similar in both data sources. It seems that retrieval rates do not cause the differences between regions, but crime rates. Either way, it is important to know that there are no large differences in the distributions of percentages of crimes committed in the different regions in the DNA that could have influenced the results of chapter 3.

Although DNA data are not free of selection bias (and in fact are quite selective because a few specific crimes such as homicide, rape and more recently volume crimes like burglary are much more likely to be include than other types of crime), this selectivity differs from the selectivity of data sources commonly used in criminological research on offenders. As Tilley and Townsley (2009) described the selectivity of DNA data differs from both traditional sources in law enforcement (for example arrest data or records from prisons) and from self-report surveys. The main advantage of DNA traces remains: they offer the opportunity to study the offending behaviour of non-arrested offenders. DNA traces thus can give other insights into offender behaviour than could be gained from arrest data or self-reported delinquent behaviour data.
5.3 Limitations of using DNA traces for criminological research

There are also more general limitations regarding the use of DNA traces for criminological research. A number of these limitations have been discussed in each of the empirical chapters in this thesis. The three limitations described in these chapters are:

1. Representativeness of the DNA data, as described extensively in the previous section (5.2).

2. The question whether collected DNA is actually offender DNA or DNA from another person. When the police visit a crime scene, they will try to make sure that the collected DNA trace is indeed DNA from the offender and not from another person present at the crime scene. They do this by reconstructing what happened and using information that is already known about the crime, and they will eliminate the victim(s) or witnesses as donors of the DNA (Meulenbroek, 2008). It could still be that a secured DNA trace is not from the offender. However, as crime series are studied in this thesis, the probability that all traces in a series are from the same person who is not the offender, is (very) small.

3. The reliability of a match between DNA traces, and between DNA traces and a person profile. Every DNA profile is extremely rare, but it is impossible to know with 100
per cent certainty whether a DNA profile is unique. Therefore there is always a small probability that a match between two crime scene traces (or a match between a crime scene trace and a person’s profile) is a coincidence. Also, errors can be made by laboratories, for example interchanging the DNA profile of the suspect with that of the victim (Van Koppen & Elffers, 2006). The Netherlands Forensic Institute (NFI) calculated that the probability of a match between two traces being a coincidence (when using a complete DNA profile) is always smaller than 1 in 1 billion (Meulenbroek, 2008). Although there is a probability that a match between two crime scene traces is a coincidence, one has to keep in mind that such a false positive has very different implications in court than in research. In court, an error margin of 1 per cent may be unacceptable, whereas in the social sciences an error margin of 1 per cent is negligible.

There are also other limitations of using DNA traces for criminological research that have not been described before in this thesis: not being able to study offender characteristics, certain crime types that are not present in a DNA database and the possible forensic awareness of offenders. These will be discussed below.

**Offender characteristics**

As the studies in this thesis show, the major advantage of DNA traces is that offending behaviour of non-arrested offenders can be studied. This is however the only aspect of non-arrested offenders that can be studied using DNA traces thus far. Because the offender is unknown, characteristics of the offender himself cannot be studied. Examples of these characteristics are age, race, gender, home location (or anchor point location) and demographic variables. As a consequence, it is not possible to study well known criminological concepts like age of onset, the age-crime curve and the journey-to-crime of non-arrested offenders.

However, from DNA itself some features of the person can be extracted, for example the gender of the person. Recently it also became possible to predict “externally visible characteristics”, such as eye-color and bio-geographic ancestry (at the level of broad geographic regions like continents) from a DNA sample (see for example Enserink, 2011; Kayser & de Knijff, 2011; Keating et al., 2012; Liu et al., 2009; Royal et al., 2010; Walsh et al., 2010; Walsh et al., 2011; Walsh et al., 2012). These characteristics may offer interesting opportunities for criminological research, for example the question: do non-arrested female offenders have a different geographical range of operation than non-arrested male offenders? could be answered using this type of information extracted from DNA. However, there is still a lot of debate about the judicial and ethical issues surrounding this type of DNA research. Also, the number of possible characteristics that can be studied is quite limited.
Types of crime not present in the DNA database

Section 5.2 has given some insight into the nature and extent of the representativeness of DNA data. An important point to add to that discussion is that some crimes will not be present in DNA data, unless the criminal law changes. The Dutch criminal law only allows DNA to be taken from persons that are arrested for committing a crime for which the law allows custody (crimes for which a maximum prison sentence of four years or more can be imposed). As a consequence of this law, police officers and crime scene examiners will not collect DNA traces at crime scenes of crimes for which the law does not allow custody, because they will not be able to get DNA from a suspect to compare to a potential crime scene trace. Crimes for which the law does not allow custody will therefore not be present in the DNA database.

Another point is that different types of crime are unequally represented in the DNA database. Fraud or harassment (for which the law does allow custody), for example, will probably only be present to a limited extent in DNA data, while burglary, vehicle theft and sexual offenses are represented in great numbers (Tilley & Townsley, 2009). In fraud cases, for example, DNA traces often are not part of the investigation, simply because this type of crime is not suitable for DNA research. DNA traces are not left behind at a ‘crime scene’ as easily and recognizable in fraud cases as in for example rape cases. It may be even hard to determine what and where the ‘crime scene’ actually is when it comes to fraud.

Forensic awareness among offenders

As described in chapter 3, multiple selection processes take place before a DNA crime scene trace is loaded onto the DNA database. The very first step of this process is that an offender has to leave a DNA trace behind at a crime scene. This step raises multiple questions: are offenders aware that they may leave DNA traces behind? And if they are, do they take precautions to prevent that from happening? And is there a difference between the first crime that an offender commits and later crimes? Or in other words: does the offender get more or maybe less careful about leaving DNA behind as his criminal career develops?

To answer these questions it is necessary to perform offender-based research. Some research is done on the opinions and thoughts of offenders on DNA traces and DNA databases. Prainsack and Kitzberger (2009) interviewed 26 convicted offenders in Austrian prisons. These offenders named TV series as an important source of information about DNA, although they do not fully trust this source. Another source of information on DNA for these offenders is conversations with other inmates. (Prainsack & Kitzberger, 2009). Clear is that offenders are aware of the existence of DNA, and that it could lead to their arrest. According to the prisoners Prainsack and Kitzberger (2009) interviewed, DNA is not considered specifically when a cost-benefit analysis is made whether or not to commit
a crime. Does this also mean that offenders do not take precautions to prevent leaving DNA behind at a crime scene? In a study among 31 Portuguese prisoners, conducted through in-depth interviews, Machado, Silva, and Cunha (2012) describe that offenders find it really hard to avoid leaving DNA traces behind at crime scenes, especially compared to fingerprints. The prisoners in the study of Prainsack and Kitzberger (2009) describe feelings of “powerlessness” and “a lack of control” (p. 52) when it comes to DNA. In contrast to fingerprints, offenders feel that they have no control over leaving DNA traces behind at crime scenes.

One study specifically focuses on the forensic awareness of offenders, and the consequences this awareness has for trying to avoid leaving DNA traces behind. Beauregard and Bouchard (2010) describe forensic awareness as “taking additional steps and adapting the modus operandi used in a crime to hide evidence in order to ultimately avoid apprehension” (p. 1161). In their research they study 72 sex offenders who together committed 361 offenses, using both in-depth semi-structured interviews and content analysis of police investigation reports. They conclude that their sample does hardly exhibit DNA removal behaviours, and that little attention was paid by the offenders to DNA left at the crime scene through the presence of semen (Beauregard & Bouchard, 2010).

It can be argued that persistent offenders may have different perceptions and maybe even different, and/or more avoidance strategies when it comes to leaving DNA traces behind at crime scenes, than first-time offenders. According to Park, Schlesinger, Pinizzotto, and Davis (2008) experienced offenders will be more likely to show forensic awareness by removing evidence from the crime scene that may be used in the investigations of law enforcement agencies. In their study on 22 serial and 22 one-time rapists, they conclude that serial rapists do show a higher level of “criminal sophistication” than one-time offenders, including more forensic awareness (Park et al., 2008). In line with Park et al. (2008), Prainsack and Kitzberger (2009) also state that first-time offenders do not tend to have expert knowledge on how to avoid traces when committing a crime. Also, as Tilley and Townsley (2009) state: “none but the most reckless offender leaves traces at all their crime scenes”.

So although offenders are aware that they may leave DNA traces behind at a crime scenes, they do not think that they can prevent it, and also do not take specific measures to prevent leaving DNA traces behind. However, experienced offenders are more forensically aware than first-time offenders and offenders will probably not leave DNA traces behind at every crime scene.

All the offenders interviewed in the above mentioned studies are offenders who got caught (either through DNA evidence or through different evidence). It may be that offenders who did not get caught have a different attitude towards DNA, and may be more forensically aware and better in avoiding leaving DNA traces behind.
General conclusion and discussion

Another point regarding the forensic awareness of offenders is the possibility of planting DNA traces. Offenders themselves may plant DNA traces of other people at a crime scene (for example by emptying an ashtray from a café at the crime scene), but another possibility is the tampering with the crime scene by police investigators or other persons by deliberately planting somebodies DNA traces on the crime scene (Griffith & Roth, 2006). Bolden (2011) for example describes that prisoners are exchanging information about how to spread blood and semen samples from other people around crime scenes.

Based on the studies described here on the forensic awareness of offenders it is hard to draw any conclusions on the selectivity of (Dutch) DNA data stemming from this possible forensic awareness. However, the results of the described studies do indicate that it is important to include the offenders in studies on the selectivity of DNA data, besides the behaviour of the uniformed police and the forensic department.

5.4 Future research possibilities

Other aspects of offending behaviour
This thesis focuses on two aspects of offending behaviour: criminal career characteristics and spatial characteristics. As explained throughout the thesis, it is not possible to study characteristics of the offenders themselves (although in the future there may be possibilities to use for example the gender of the offender). It may however be interesting for future research to focus on different aspects of offending behaviour. An example of this is the time on which the crime was committed, for example the different days of the week (offenders who only commit crimes in weekends vs. offenders who commit crimes during weekdays), the time of day, or the time of year, or the time between offenses. Another option could be studying (aspects of) the modus operandi. Both the time and the modus operandi have to be retrieved from police data, because the DNA database does not contain enough detailed information on these topics. Consequently, these data may have similar problems with missing values as the data used in chapter 4. Still these may be interesting aspects of offending behaviour to study using DNA traces, so that a comparison between arrested and non-arrested offenders can be made.

A last aspect that could be studied is offenders who commit only one single crime. The data used for the current study also contained a large number of single crimes. Because this thesis focuses on crime series, these are not taken into account in the analysis. However, they could be used to study for example whether different crime types have different clearance rates and whether this is different in different police regions.

Offenders who operate internationally
It may be interesting to use the DNA database to conduct studies on offenders who operate internationally. What influences the probability that offenders who operate internationally
Chapter 5

Will be arrested? Currently, Dutch DNA profiles and crime scene traces are compared for investigative purposes on a daily basis to both DNA profiles and crime scenes traces from DNA databases of Germany, Austria, Spain, Luxembourg, Slovenia, Finland, France, Bulgaria, Slovakia, Romania, Latvia, Lithuania and Hungary. However, no information about the crimes itself is automatically exchanged. If there is a match between crime scene traces or between a crime scene trace and a person profile and more information about the crime or the person is needed, and official international legal request needs to be submitted before this information can be exchanged (NFI, 2009, 2013). A permanent exchange between DNA databases in different countries for research purposes is thus not (yet) possible.

Offender based research
Other interesting future research could be offender based. Was there anything different about the crime the offender was arrested for, compared with the crimes for which he was not caught? In other words, does the offender think that something in his behaviour may have influenced the probability of getting arrested? Other interesting offender-based studies may focus on the forensic awareness of offenders, to expand the limited number of studies that are done on this topic. Offenders that have not been arrested can then also be taken into account, to study whether they have different opinions on DNA than the incarcerated offenders studied thus far. An interesting question for example is whether non-arrested offenders think they are able to avoid leaving DNA traces behind. This offender based research on non-arrested or non-incarcerated offenders for example can be done by using snowball sampling. See for example the studies by Jacobs, Topalli, and Wright (2003), Jacques and Reynald (2012), Jacques and Wright (2011), and Mullins, Wright, and Jacobs (2004).

Capture-recapture
Traces series could also be used with the capture-recapture method to estimate the size of the entire offender population (see for example on the criminological use of the capture-recapture method Bouchard, 2007; Rossmo & Routledge, 1990; Van der Heijden, Cruyff, & Van Houwelingen, 2003). Normally when this method is used in criminology, the (repeated) arrest of an offender is the starting point of the analysis, but in the case of DNA traces it would be the (repeated) discovery of a DNA trace from the same offender.

Influence of the DNA database on deterrence
Future research may also focus on how the fact that an offender’s DNA profile is present in a DNA database influences the probability of clearance. In this thesis, crimes that had been committed after the offender was arrested and after a DNA sample had been taken from them and loaded onto the database are not taken into account. Does the storage of a
DNA profile in a national DNA database deter offenders? In an analysis of data from the UK National DNA database (NDNADB), Leary and Pease (2003) show that there is no increase over time in the proportion of crime scene samples that match an offender profile. This indicates that the pool of active offenders is unstable: many of the offenders for whom DNA profiles are stored in the NDNADB no longer offend and many of the offenders who do offend have not yet been included in the NDNADB. Future studies may expand on this type of research.

5.5 Policy implications

This last section of the discussion discusses what the consequences of the results of this thesis are for both criminological studies and policy decisions that are based on arrest data.

Using arrest data for criminological research

Criminal career research

Many criminologists who study criminal careers use arrest or conviction data to do so (see for example Francis, Soothill, & Fligelstone, 2004; Mazerolle, Brame, Paternoster, Piquero, & Dean, 2000; McGloin, Sullivan, Piquero, & Pratt, 2007; Miethe, Olson, & Mitchell, 2006; Nieuwbeerta, Blokland, Piquero, & Sweeten, 2011; Piquero, Brame, Mazerolle, & Haapanen, 2002). A number of authors have expressed their concerns about using arrest data to study criminal careers, because only a small percentage of offenders will be arrested, and these offenders may not be representative for the entire offender population. Piquero, Hawkins, and Kazemian (2012) for example state that specialization is primarily studied using arrest and/or conviction records and that these studies may have captured only more serious crimes because of the bias in these types of data. Lynam, Piquero, and Moffitt (2004) also describe this probable bias of arrest data: “Official data not only underrepresent the degree of total offending but may over-represent more serious, violent crimes that are cleared at higher rates”. To get insight into and control for these problems of official data, some studies compare arrest data to self-reported data, or use both arrest data and self-reported data to validate their results (see for example Babinski, Hartsough, and Lambert, 2001; Brame et al., 2004; Kirk, 2006; Lynam et al., 2004). According to Brame et al. (2004) it is important to use both official records and self-reported data to study the same individuals as both data sources have their particular strengths and weaknesses. Piquero et al. (2012) find that because most offenders are not caught, self-reported involvement in criminal behaviours tends to show much larger numbers on involvement in crime then official records from the police. Lynam et al. (2004) in their study on offense specialization find that in official reports (in this study: conviction records) no evidence for specialization is found, whereas in self-reported data specialization was evident. They say that the results from the conviction data records are in line with results from previous studies on sequential specialization. Kirk (2006) draws a
similar conclusion: “Findings reveal that a sizable number of youth self-report being arrested without having a corresponding official arrest record, and a sizable proportion of those youth with an official arrest record fail to self-report that they had been arrested.” In contrast to these results, Brame et al. (2004) find that although not many studies use both data sources to study the criminal behaviour of the same individuals, in the studies that do use both sources there is a modest positive correlation between self-reported and official record measures of involvement in criminal behaviour. And although they themselves do not do this, Mazerolle, Piquero, and Brame (2010) think that it is important to replicate their study (for which they used correctional records) with self-reported data.

Result from this thesis can add to this discussion, as they give insight into the representativeness of arrest data, as explained before. Results from chapter 2 show that a number of criminal career characteristics have an influence on the probability that an offender will be arrested. Interesting is that one result from this thesis is that specialized offenders have a smaller probability of being arrested than versatile offenders, which means that specialized offenders are under-represented in arrest data. This is in accordance with result from Lynam et al. (2004) who find no evidence for specialization in the conviction records they used for their study, but do find specialization in self-reported crimes. The results from the current study suggest that arrest data are not representative when they are used to study offense specialization.

Another finding from this thesis is that as the number of crimes an offender commits increases so does the probability that he will be arrested. All the analyses in each of the chapters that deal with answering the main question show this result. More prolific offenders may therefore be overrepresented in arrest data. A last result from this thesis concerning criminal careers is that seriousness does not have an influence on the probability of arrest, so the concerns of Piquero et al. (2012) and Lynam et al. (2004) that more serious crimes are over represented in arrest data may not be justified.

Research on spatial offending behaviour
Chapter 4 describes a discussion of a similar nature about the use of arrest data to study the spatial behaviour of offenders. From that chapter it became clear that the average distance between crime locations in a series of offenses does not influence the probability that an offender will be arrested. This means that offenders present in arrest data are representative when it comes to this specific measure of spatial behaviour. Chapter 3 also studies an aspect of the spatial behaviour of offenders: the number of police regions in which crimes are committed. If the number of regions increases, the probability of arrest decreases. This may indicate that offenders with a greater working area are under-represented in arrest data. However, this is contradictory to the results of chapter 4 in which an increase in distance does not cause an increase in probability of arrest. At the end of chapter 4 the probable causes for these inconsistent results were already discussed. It
may be that the cooperation between police regions is not optimal, and that this causes the decreased probability of arrest, and not the spatial offending behaviour itself. For the selectivity of arrest data this means that when the spatial behaviour of offenders is studied, still some caution is necessary, because it is not entirely clear whether or not arrested offenders are representative for the entire offender population. It is however clear that in term of the average distance between crime locations, arrested offenders do not differ from non-arrested offenders.

Concluding it can be said that the results of this thesis show that arrested offenders are not entirely representative for the entire offender population, although some concerns (about more serious crimes being over-represented in arrest data) might not be entirely justified. Clear, however, is that persistent offenders are over-represented in arrest data, relative to the proportion of persistent offenders in the offender population. Less problematic is the use of arrest data to study the spatial dispersion of the crime locations of an offender, because offenders present in arrest data are representative in terms of the average distance between their crime locations. Researchers and policy makers that use arrest data for their study of offenders should take the results of this thesis into account.

**Information sharing between police regions**

Another tentative policy implication from this thesis may be that information sharing and cooperation between different police regions could maybe be more effective. The conclusion from chapter 3 suggests that more cooperation and information sharing between different police regions may increase the number of arrested offenders. Since January 2013 the Netherlands has a new police system, which consists of one national police force and ten regional forces. The practical implementation of this new system may take a couple of years, however it would be interesting to study in 10 or 15 years whether this new system has helped the cooperation between regions, in terms of an increase in the number of arrested offenders. Almost every region from the previous 25 regions form a new region with another former region (see Figure 5.2 a and b).
Figure 5.2a  25 original police regions  Figure 5.2b  10 new police regions
References


