Immigration Status and Victimization

Evidence from the British Crime Survey

Georgios Papadopoulos*

School of Economics, University of East Anglia

March 2013

Abstract

This study, using data from the British Crime Survey (BCS), examines the micro-relationship between immigration and victimization. We first find that, although immigrants are more likely to suffer property crimes than natives, this is well explained by the fact that immigrants exhibit demographic characteristics associated with higher victimization. Contrary to the above, immigrants are of lower risk of violent victimization. As violence is an expressive type of crime, where interactions between victim-offender pairs prior to the incident matter more than instrumental crime, the lower risk of violence can be attributed to different lifestyle choices associated with lower victimization risks. However, a closer investigation, decomposing violence in domestic, by acquaintances and by strangers crime, shows that this difference is driven by the lower crime immigrants suffer by acquaintances and by family members, which is not consistent with the 'different-lifestyles' hypothesis. Nevertheless, we show that the aforementioned (unexpected) difference cannot be attributed to higher under-reporting by immigrants. We further show, that if immigrants did not face racially motivated crime, they would also face a significantly lower risk of victimization by strangers. Finally, we examine whether the lower victimization by acquaintances could be because more recent immigrants have fewer acquaintances. However, we argue that if this kind of "network" effect exists, it is actually quite weak. Therefore, all evidence suggests that indeed, immigrants face a lower risk of violent victimization because of lifestyles associated with a lower exposure to crime. Finally, using count data models we examine whether immigrants are disproportionately victims of repeat crimes. However, the results show that patterns of repeat victimization are generally the same between immigrants and natives.

Keywords: Crime; Victimization, Immigration, Self-completions, Reporting Behavior, Repeat Victimization.

JEL Classification Numbers: K42, J15, C18, C21, C24, C25

---

*Address: School of Economics, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. E-mail: *G.Papadopoulos@uea.ac.uk*.
I am grateful to João Santos Silva, Alison Booth, Zelda Brutti, Michail Veliziotis, Mariña Fernádez Salgado, and participants in the EEA 2011, ESEM 2012, RES 2012 conferences and the Crime Surveys User Meeting, ESDS, London, for their helpful comments. The empirical analysis is implemented in Stata/SE® and TSP® econometric software. The data used in this project are sponsored by the Home Office and provided by the UK Data Archive. Neither the original data collectors nor the aforementioned individuals bear any responsibility for the analyses or conclusions presented here.

1 Introduction

The link between immigration and crime is well discussed among scholar and non-scholar communities. Nevertheless, most of the discussions by non-scholars concern immigrants' involvement into criminal activities as offenders. This one-dimensional treatment has led to the sentiment that immigrants are more involved in illegal actions. This is in most cases in contradiction with scholars' findings which often suggest the opposite.[1] For instance, Papadopoulos' (2013) findings, in a study for England and Wales, suggest that immigrants' participation in criminal activities as offenders (both in property and violent crimes) is slightly lower than natives' one as opposed to the public sentiment.

Since scholars have also focused on the immigration-crime link from the offending point of view, they have overlooked the other important side of the coin which concerns the engagement of immigrants in crime as victims. To my knowledge there are no comprehensive studies that concentrate on this relationship,[2] as most studies focus on the determinants of victimization[3] in general (see, for example, Miethe, Stafford and Long, 1987, Smith and Jarjoura, 1989, Wiles, Simmons and Pease, 2003, and Tseloni, 2006). More relevant studies look at the experiences of ethnic minority groups without distinguishing between native and immigrant populations (see, for instance, Clancy at al, 2001, Jansson, 2006) and some of them are criminological studies that focus on a very specific aspect of victimization experiences by ethnic minorities, namely, racially motivated crime, or differently, 'hate' crime (see, for example, Gabbidon and Greene 2009, and Kalunta-Crumpton, 2010).

However, the two sides of crime, namely, offending and victimization, cannot be considered as independent actions, especially for violent crime. Actually, being an offender signifies a social lifestyle associated with a higher likelihood of victimization compared to a non-offender (see, Lauritsen, Sampson and Laub, 1991, and Deadman and MacDonald, 2004). This is recognized in very early theories of criminology (see, for example, von Hentig, 1940) and incorporated into the succeeding, more comprehensive, theories of *lifestyle-exposure and routine activities* (see, Cohen and Felson, 1979). Therefore, looking at the association between being an immigrant and victimization may reveal some interesting insights on the general immigrants' behavior towards criminal activities relative to natives' one. This study, therefore, intends to fill the gap in the literature and complete the immigration-crime picture by investigating the victimization differences between immigrants and natives in England and Wales.[4] Moreover, it will also shed some light on the social integration of immigrants. As a result, the findings of this work could

[1]For a review of the literature for the involvement of immigrants as offenders refer to Papadopoulos (2013)

[2]To the authorÕs knowledge, the only finding on this link for the UK comes from Bell, Machin and Fasani (2010) study, who find that immigrants are less likely to be victimized using British Crime Survey data from 2004 to 2008. However, experiences in terms of victimization concerns only a very small part of their paper, so that they give only a very narrow picture of the immigration-victimization link.

[3]It is important to stress at this point that in this paper, victimization is defined as the incident of becoming a victim of a crime. We stress this in order to avoid confusions, since the term victimization is nowadays also interpreted as the act of exploiting or treating unfairly someone.

[4]This study examines the victimization experiences of immigrants and natives only in England and Wales, as Northern Ireland and Scotland are excluded from the British Crime Survey because of their distinct criminal justice system.

also be a useful tool for policy makers.

The main aim of the present study is to comprehensively examine whether immigrants are more or less likely to become victims of crime and whether differences would still exist if immigrants shared the same demographical characteristics with the native population. For the purposes of the above analysis we use the 2007-08 sweep of the British Crime Survey (BCS), a representative victimization survey where respondents were asked in face-to-face interviews about their victimization experiences in household and personal crime. We need to note that the nature of the victimization incident is very different across different crime categories, such as property crime (burglaries, vehicle thefts, other thefts, criminal damage) and personal crime (personal thefts, violence). Therefore, the immigration-victimization link will be examined separately for the different crime categories, but more attention will be paid to violent crime. An important point is that, since the questionnaire of the BCS involves some questions that try to elicit very sensitive information, misreporting is a concern. For instance, there is evidence that respondents tend to under-report domestic violence perhaps because of fear of reprisal, or because they want to protect the offender (Walby and Allen, 2004, and Felson at al, 2006). If immigrants' reporting behavior differs from natives' one, the estimated coefficient representing the difference in domestic victimization between immigrants and natives will be biased. However, in Section 6, by we show that immigrants do not under-report by more than natives.

According to theories of victimization (see next section), there are many channels through which immigration and victimization are linked either positively or negatively. For instance, immigrants are more likely to be victimized just because they are disproportionately located in deprived areas where crime rates are much higher. On the other hand, immigrants are less attractive as targets since they usually possess fewer properties, or relatively less valuable objects. Therefore, the theory cannot provide a clear-cut relationship between immigration and pecuniary, or differently, pecuniary crime;[5] this is rather an empirical question. As our data provide a lot of information on characteristics that are associated with instrumental crime, we are able to acquire a better understanding of the reasons why we (do not) observe different risks of victimization between immigrants and natives.

The case of violent crimes is less obvious, however, as violence refers to expressive actions where the offender intends to hurt the person and not to acquire his/her property. In violent crime, contrary to property crimes, interrelations and interactions among potential offenders and potential victims are important. Thus, personal behavior is a much stronger predictor of violent victimization compared to instrumental victimization. Although theories of victimization are still valid (given some conceptual modifications), it is difficult to identify the theoretical channels through which immigration status in associated with higher or lower violent victimization, as most determinants of violence are unobserved factors determined by the potential victim and his/her relationship and interactions with potential offenders. For instance, some people would be less likely to suffer a violent crime if they followed particular lifestyles associated with lower

[5] By pecuniary/instrumental crime we mean a crime where the offender's intention is to acquire victim's property and not to hurt the person itself.

crime. However, since most aspects of this lifestyle are unobserved, only speculations can be done to explain the factors that have generated this differential risk of victimization among different groups of individuals. Moreover, violence consists of three crime types of very different nature, namely, crime suffered by strangers, crime suffered by acquaintances and crime suffered by family members. As will be seen later, modeling these three crime types separately will provide some very interesting insights on the immigration-victimization nexus.

Once the above relationships are established using a thorough examination of sensitivity tests, we exploit the number of victimization incidents using appropriate regression models for count data to develop a better understanding of immigrants' victimization experiences relative to natives' ones. Differences in the estimated coefficients between the binary models (victim or not) and count data models will indicate that there are differences in repeat victimization experiences between immigrants and natives. Special attention will be paid to the substantial concentration on the zero outcome (no crimes suffered) and to the presence of some extreme cases where people reported an extreme number of suffered crimes. The models used in this paper recognize that the process generating the zero-one outcomes might be different from the one generating the positives, which implies that the effects of the explanatory variables may differ between the probability of a single incident and the number of incidents conditional on having already suffered a victimization incident. Other interesting topics will also be investigated, such as, whether the ethnic composition or the location of immigrants matters, whether there are assimilation patterns in the immigrants' victimization experiences and whether immigrant victims perceive their victimization experiences as more serious than otherwise comparable natives.

The rest of this study is organized as follows. In the next section a brief exposition of a victimization theory together with a short discussion of the link between immigration and crime is presented. Section 3 is devoted to explaining some technical parts of the BCS and the construction of the dependent variables. Additionally, a description of the data used in the empirical analysis and some descriptive statistics are presented. In Section 4 a basic analysis for household crime follows, where we investigate whether immigrants are more or less likely to be victims of household crimes, focusing on inside and outside burglaries. Section 5 examines the experiences of personal crime. Although some results on personal theft are also presented, this section concentrates on violence. Section 6 provides a thorough sensitivity analysis with regard to the results between immigration and and the risk of violent victimization. Section 7 presents a few results of interaction terms and perceived seriousness of victimization incidents. A comprehensive analysis of count data models follows in Section 8. Finally, Section 9 consists of concluding remarks.

2 Theoretical Perspectives on Victimization

Before discussing the theoretical concepts of victimization it is worth noting that although this study also presents results on inside and outside burglaries and on personal thefts, most of the attention is paid to violent victimization. The results of other crime types, such as vehicle crime, criminal damage and household thefts will be briefly discussed but not presented in detail. Moreover, as these crime types are of a very different nature, we need to emphasize that to some extent the theoretical concepts apply differently to the different crime groups.

When it comes to the offender-victim relationship it is natural to argue that in many cases full responsibility falls onto the offender (although victims could still be unintentionally responsible). For instance, think of a young girl whose purse gets stolen in a busy train station. This is not always the case though. Even early theories (see, for example, Von Hentig, 1940 and Wolfgang, 1958) admit that there are cases in which offenders do not bare full responsibility, but the crime is a function of the underlying offender-victim relationship evolving prior to the victimization incidence. Crimes are considered as interactive acts that depend upon the actions of both parts. Thus, these theories rule out the factor of "randomness" in victimization incidents.[6] For instance, the theory of precipitation, first discussed by Wolfgang (1958), argues that to some extent it is the victims' provocative behavior that initiates subsequent crimes against them (see also, Schultz, 1968, and Curtis, 1974). Clearly, the above theories seem more appropriate to describe violent crimes than instrumental crimes where for instance, the victim using gestures or offensive language initiates an assault. Or, we could think of a case of domestic crime where the interaction of family members is very important.

However, the theories that have attracted most both theoretical and empirical research are based on the concepts of *lifestyle-exposure* (Hindelang, Gottfredson, and Garofalo, 1978) and *routine activities* (Cohen, and Felson, 1979). Earlier concepts, such as the importance of offender-victim relationship are integrated into these ones. We need to note that although each of these two theories was initially developed for different purposes, they are closely related and the present study treats them as a single comprehensive theoretical framework (see, Meier and Miethe, 1993, for an elaborate exposition of these theories). According to them, *routine activities* and particular *lifestyles* of potential victims shape a criminal opportunity structure which consists of four distinct risk factors that are associated with victimization. These factors are: *proximity*, *exposure*, *attractiveness* and *capable guardianship*. *Proximity* and *exposure* create the criminal opportunity structure, whereas *attractiveness* and ability of effective *guardianship* determine the criminals' choice of victims (Miethe and Meier, 1990).

Proximity is defined as the physical distance between the location of potential targets and the location where potential offenders mostly act. For instance, living in highly deprived areas, where the crime rates are high, increases the probability to be victimized, as it increases the probability of contacting potential offenders. This concept becomes less relevant as the mobility of the target increases, since the task of identifying the distance between offenders and victims becomes more difficult. Therefore, although the concept of *proximity* is very clear for household crimes, it loses some transparency once we deal with personal crime. However, it is still important as most victims tend to socialize in areas close to their residences.[7]

[6] By "random" victimization incidents I mean situations where, there is no prior relationship between the offender and the victim and the victimization incident does not depend on the interaction between offenders and victims.

[7] According to the victim forms of the 2007/08 BCS around 20% of all personal victimization incidents happened

Exposure refers to the physical visibility or availability of potential victims. The meaning of this concept changes substantially across different types of crime. For personal violence, *exposure* can be conceptualized as the general *routine activities* or *lifestyles* of potential victims, associated with higher or lower likelihood of victimization. For instance, people that mostly stay at home and do not socialize in bars or pubs tend to be less likely to suffer a violent crime. Here, general lifestyle also includes relationships and interactions of potential offenders with potential targets. Thus, this concept also incorporates the earlier theories of precipitation. However, for household crimes, for instance for inside or outside burglaries, *exposure* may refer to the location of the house (such as main road or cul-de-sac), or the amount of properties someone possesses. For vehicle crime just a high number of cars owned by an individual can be considered as an indicator of high *exposure*.

Target attractiveness is defined as the material (for acquisitive crimes) or symbolic (for violent crimes) desirability (value) of targets to potential offenders. The notion of *attractiveness* is again very different across acquisitive and violent crimes. For instance, in household crime of acquisitive nature, the appearance of the house, or the information of offenders for valuable objects inside the house increases *attractiveness*. For personal thefts, the general appearance can indicate a level of *attractiveness*. On the other hand, violence is an expressive crime, as offenders target to hurting the victim itself without being interested in victim's valuable possessions. [8] Just the ethnicity of a potential victim can be considered as a highly attractive attribute for an extremist. In other cases *attractiveness* develops through interactions and interrelations among people. For example, a member of a gang finds as an attractive target a member of another gang (with regard to the symbolic utility that the offender gains if they commit the crime).

Finally, *physical* or *social guardianship* is the effectiveness of objects (*physical guardianship*) or people (*social guardianship*) in preventing crime from occurring. For personal crimes, *guardianship* is the ability of the person, or the ability of people around them, to protect himself/herself. Having a weapon in apparent place, or guards, is a type of *physical guardianship*. Also demographic features as height, weight, age, appearance, could indicate an ability of protection. *Physical guardianship* for dwellings and vehicles could be for example security measures, neighborhood watching program, etc. On the other hand, social measures could be the number of hours a house is left unoccupied, number of household members (more members indicates that the house is left unoccupied less hours per day, which decreases the likelihood to be burglarized), knowledge on what to do in case someone breaks into the house, etc.

The basic economic theory of crime is closely related to the above sociological views. A two-stage model which borrows simple notions from the early economic models of crime by Becker (1968) and Ehrlich (1973) could be formulated to describe the victim-offender relationship. According to these early models of crime, individuals use a rational cost-benefit analysis where they

inside or immediately outside victims' residence. From the rest of them, 6% occurred in workplace, 18% at pub/bar/club, 35% in other public or commercial location and 22% elsewhere. Moreover, it is very interesting that for the incidents that did not happen inside or outside residence, 40% of them took place within 15 minutes from victim's residence.

[8]It is important to note that for robberies there is a violent act together with the theft. However, as the primary target of the offender is instrumental, I consider robbery as personal theft.

weigh the expected costs and benefits in utility terms and subsequently decide how much time to allocate in legal and criminal activities in order to maximize their net expected utility. Since crime involves uncertainty, because of potential apprehension and consequent future punishment, the concept risk aversion is very important. At the same time, uncertainty and risk aversion are also very important from the potential victims' point of view, as the actions of potential victims could not perfectly determine the criminal activity against them.[9]

Although in reality the situation is much more complicated, a simple model could be formulated as follows: in the first period the (rational) potential victims, given the level of risk aversion and initial values of *attractiveness*, *exposure*, *proximity* and *capable guardianship* (as these are determined by their exogenous socio-economic and demographic attributes), consider a set of different strategies and the possible subsequent actions of potential offenders for each different strategy. Consequently, they re-evaluate their position by determining to some extent the optimal levels of *attractiveness*, *exposure*, *proximity* and *capable guardianship* in order to maximize the net benefits. For instance, people that are highly afraid of potential offenses (such as older people), which denotes high risk aversion, would decide to exhibit very low *exposure*, for example, by staying mostly at home and avoiding going out at night, or to increase *guardianship* by taking higher physical measures of protection. On the other hand, people that value enjoyment by much more than safety (such as younger people) would disregard many potential dangers and exhibit high *exposure* and *attractiveness* for the sake of amusement.

In the second stage, once the opportunity criminal structure is set by the determination of *proximity*, *exposure*, *attractiveness* and *guardianship*, potential criminals come into play. Each of the four risk factors can be translated into costs and benefits for the offender. For instance, a highly attractive person or household would result in higher utility for the offender, a well protected house increases the uncertainty of success of the criminal action and therefore increases costs, a household of high *exposure* decreases uncertainty and therefore, decreases costs, and so on. Consequently, potential criminals, comparing their legal and illegal opportunities and taking into account their criminal ability and risks they are willing to take, decide whether to commit crimes and consequently which targets to hit in order to maximize their expected utility. Of course, the whole procedure is more complicated since potential victims cannot perfectly observe the actual risks of victimization for each strategy they follow, and in a similar manner in the second period the four risk factors are not perfectly observed by the potential criminal. Moreover, this model also ignores the possibility that potential victims can at the same time be potential offenders. Nevertheless, this simple setting together with the socio-criminological views could give some predictions on the immigrant-victimization relationship.

We need to emphasize that all ascribed or acquired attributes, such as age, gender and race, or education, income, marital and employment status, respectively, are associated with victimization likelihood through their effects on the described risk factors. For example, males generally prefer to socialize more frequently in dangerous places and they exhibit a more aggressive behavior relative to females. Therefore, they would decide to be more exposed to criminal activities,

[9]For instance, a burglary cannot be avoided with certainty even if the potential victim is very cautious.

which makes them more likely to become victims of violence. However, the situation is very different for domestic crime. Males within a family are victimized to a lesser degree because they exhibit higher *guardianship*. Moreover, the effect of some other attributes is ambiguous as they affect victimization risk through two or more risk factors. For instance, more affluent households are associated with both higher or lower risk of a burglary, since high household income may indicate a better protected house (more *capable guardianship*) or a very attractive target.

2.1 The Immigration-Victimization Link

Although immigrant population is rather heterogeneous, the 'typical' immigrant shares some common characteristics. In Table 3 some descriptive statistics from the BCS 2007-08, by immigration status, can be found. From this table it is clear that immigrants are relatively younger and more from ethnic minorities, they are more unemployed and disproportionately located in deprived inner city areas, mostly of London, and there is also evidence that they are on average poorer and face lower legal opportunities relative to natives (see, for example, Algan at al, 2010). Given all the above characteristics, and assuming that labour outcomes enter the problem exogenously, immigrants evaluate their initial levels of *attractiveness*, *exposure*, *proximity* and *guardianship*, as all these socioeconomic characteristics are to some extent associated with these four risk factors. Consequently, they reevaluate their position by following strategies that minimize the victimization risks for each crime group given all the aforementioned constraints.

For instance, location of immigrants, and consequently *proximity*, is constrained by their labour outcomes. As immigrants face unfavorable labour outcomes they can only afford to reside in areas where the rents are relatively low. It happens that these areas are relatively more deprived with high crime rates and therefore, of higher *proximity*. Nevertheless, given the above constraint, immigrants reduce the risk of both personal and household victimization by choosing to reside (within these areas of high *proximity*) in locations with high concentration of the same ethnic group. This strategy provides a higher insurance against risk of victimization by increasing *social guardianship*.

Moreover, as immigrants disproportionately belong to ethnic minority groups they are in higher danger of racially motivated violence, for example by being relatively more attractive to extremist groups. Therefore, they might choose to balance this unfavorable position by choosing *routine activities* and lifestyle *exposure* associated with lower victimization. In addition, a proportion of immigrants might feel alienated and react in this perceived hostile environment by following strategies that reduces the risk of victimization. Finally, immigrants could naturally exhibit different *exposure*, because of cultural differences that are associated with different lifestyles.

As mentioned in the introduction, violence consists of three crime types of very different nature, namely *crime by strangers*, *crime by acquaintances*, and *domestic crime*. Theoretical predictions on the association between immigration status and domestic crime or crime by acquaintances can be given by immigrants' relative participation in the illegal sector as offenders.

For instance, an immigrant's pool of family members and acquaintances is likely to include a high percentage of other immigrants. If we accept that immigrants, according to Papadopoulos (2013), are slightly less likely to commit violent crimes, holding everything else constant, we would expect a negative relationship between being an immigrant and violent crime suffered by acquaintances or family members.[10] On the contrary, since the proportion of native strangers that immigrants meet should not be very different from the proportion of native strangers that natives meet, the above idea does not apply.

A negative relationship could also be observed because of 'network effects', a concept closely related to *exposure*. For instance, *ceteris paribus*, immigrants will be less likely than natives to suffer a crime by acquaintances due to the fact that the 'pool' of acquaintances is smaller for immigrants than for natives, which is particularly true for most recent immigrants. According to this, we would expect that as time spent in the host country increases, immigrants enlarge their group of acquaintances, and therefore, to some extent they assimilate to natives' risk of victimization by acquaintances.

As it is clear from the discussion of this section, unobserved interactions and interrelations among people are relevant for violent crime, but not for household burglaries and personal thefts. Household burglaries more or less depend on observed household characteristics. The fact that the household reference person (HRP) is an immigrant should not affect the risk of victimization, given that we are able to control for all household characteristics associated with burglary victimization.[11] Fortunately, the BCS provides a rich set of household characteristics directly associated with *lifestyle-exposure* and *routine activities*, such as hours home left unoccupied, being in a neighborhood watching program, house condition, type, location, etc (see, next section).[12] The situation of personal thefts is a bit more complicated due to the fact that it entails personal contact and thus, the potential criminal can directly observe the potential victim. However, as for burglaries, personal theft is more "random" in the sense that personal behavior is not an important predictor of the action.

Nevertheless, the risk of violent victimization highly depends on the unobserved strategies associated with particular *lifestyle-exposure* and *routine activities* that immigrants and natives set in order to reduce the victimization costs. As described above, *lifestyle-exposure* and *routine activities* might be very different between immigrant and native groups and therefore, the theory cannot provide a clear-cut relationship. This should be instead established by the empirical analysis.

[10]As an example, consider the following simple calculation. Assume that the probability to commit a crime is 6% and 10% for an immigrant and a native respectively. Also, assume that 5% of natives' acquaintances are immigrants, but 60% of immigrants' acquaintances are immigrants. According to these assumptions, holding everything else constant, the probability for an immigrant to suffer a crime by an acquaintance is $6\% \times 0.60 + 10\% \times 0.40 = 7.6\%$, but this figure is 9.8% for natives, so that the difference is 2.2 percentage points.

[11]Unless criminals seek places that are inhabited by immigrants, or criminals have information about the immigration status of residents and tend to prefer targeting these places. However, for a household crime it is the instrument that is more important than the person who owns it or resides in it.

[12]The only characteristic that might be important to describe instrumental victimization risks but is not observed in the data, is the "size" of potential victims' possessions (apart from the number of vehicles which is observed). This can be considered as more important for properties outside the dwelling as they are directly observed by potential criminals, as opposed to interior properties.

3 The British Crime Survey and Descriptive Statistics

The BCS 2007-08, carried out by the Home Office, is a representative (primarily) victimization survey where respondents in England and Wales were asked in face-to-face interviews about their victimization experiences in both household and personal crime. As will be described later, the BCS also includes computer-based self-completed interviews for the more sensitive crimes, such as domestic violence and sexually motivated offenses. Moreover, it does not interview people from Scotland and Northern Ireland as they now conduct separate surveys. The reference period for all interviews refers to the victimization incidents during the last 12 months prior to the date of the interview. It is one of the largest social surveys in England and Wales as it interviews approximately 47,000 respondents per year.

This survey is ideal to identify determinants of victimization since, together with information on victimization experiences, a large set of demographic characteristics together with information on household and personal characteristics associated with victimization are available. Note that, since they interview only private households, it does not cover commercial victimization, frauds and victimless crime, crime against children, crime against people currently in institutions, and murders (for details refer to Bolling, Grant, and Donovan, 2008).

For the purposes of this study, information from three separate files from the BCS 2007-08 was combined using the unique identifier variable from the three data sets. These files are: 1) the Main BCS data set, where information for all respondents and their households is included, regardless of their victimization experiences; 2) the Victimization Form data set, in which details of each crime reported by victims are given, and 3) the Self Completion data set of domestic violence, where all people between 16-59 years old, by participating in computer-based self-reported interviews, provided information on their domestic violence experiences.

These three data sets were constructed by using a complicated procedure whose main steps are briefly described as follows: respondents, after giving some information on demographic and other individual and household characteristics, were asked a list of screener questions about whether they suffered any type of victimization incidents during the last 12 months (against them or against their household). In case the respondent reported a suffered crime, a victim form was given for each crime suffered. The victim forms assigned to victims, which were limited to a maximum of six, contains detailed information about each crime incident. This information was next used by trained coders to assign either a valid or an invalid victimization code.[13] The cases in which the conductor was uncertain about the offense code to be assigned were sent to Home Office to be crosschecked by Home Office experts. There, a finalized code was assigned. If a particular crime in a given victim form was described as a 'series' crime, where a series crime is defined as "the same thing, done under the same circumstances and probably by the same people", the number of the incidents was recorded. The classification of crime codes is depicted

[13]The incidents are given invalid codes if the offense was a duplicate, if the offender was described as mentally ill, if the offender was a police member on duty, and if incidents that initially were given a victim form decided to be coded as no crimes after a scrutinized examination. Note that incidents outside England and Wales were given a valid code.

in Table 1.

It is important to note that some incidents include a sequence of crime events which might beof different nature. For instance, we could imagine a case where a stranger breaks into a house to steal valuables but during the burglary the house owner tried to prevent the incident resulting in suffering an assault with serious wounding. Eventually, the offender also burned the house. This incident (which is of course extreme and not very likely to have happened) involves three separate crimes but it will be recorded as an arson because arson takes priority over burglaries and serious wounding. In similar cases the final coding depends on the seriousness of the incident. For details on the coding and which crimes take priority over other ones refer to Bolling, Grant and Donovan (2008).

The outcome variables used in this study were created from the offense code variable given in the Victim Forms data set (see, Table 1). As the question of interest in this study is to identify whether immigrants are more or less likely to be victims of criminal activity (and in extension whether immigrants are more frequently victimized than natives) a grouping of the individual codes was required. Otherwise there were not enough positives in the dependent variable to precisely and robustly estimate the coefficients of the models under investigation. Seven main groups were constructed according to the nature of each crime code (as judged by the author being of the same nature), five of them for household crime and two for personal crime. For household crime these are: Inside Burglaries (codes 51-53), Outside Burglaries (codes 50, 57, 58), Vehicle Thefts (codes 60-64, 71, 72), Household Thefts (codes 55, 56, 65-67, 73)[14] and Vandalism (codes 80-86).[15] Regarding personal crime, these are: Personal Theft (codes 41-45) and Personal Violence (11-13, 21).[16]

The binary information, whether the individual reported a victimization of a particular crime type, is used in the first part of the empirical analysis (Sections 4-7), while the count information, number of crimes of a particular crime type suffered, is used in the second part (Section 8). The count variables are created by using the 'series' information from the victim forms. For example, if an assault with wounding was considered as a 'series' crime, the number of assaults forms the count variable. Moreover, if the same individual suffered another assault, for instance without injury, the number of assaults from this victim form, as indicated by the 'series' information, are added to the previous count.[17] Finally, note that it is possible two victimization forms to be assigned by the conductor for two very similar crimes, which even belong to the same code, if some characteristics of the first (series of) incident(s) are considered by the coders to be different

[14]Separation between inside, outside and other thefts was also considered.

[15]A separation between home vandalism and vehicle vandalism was also considered.

[16]For details on the crimes that each individual code included refer to the Offence Coding Coders Manual in Bolling, Grant, and Donovan (2008, II).

[17]We need to note that the main data set provides derived crime variables which are used by the Home Office to calculate prevalence and incidence rates. However, for each crime code in these variables a cap of five crimes is imposed. Therefore, the total count for a crime group will be the sum of crimes from each victimization form that fall within this crime group, where the number of crimes in each victim form is censored in five crimes. Thus, the resulting count variable will be the sum of up to six censored at five crimes. According to this, it is not proper to use a simple right censored at 30 crimes count data model but a model that allows for censoring at 5 crimes for each victimization form someone gets. This of course will result in a very complicated situation. Moreover, these derived variables do not include cases where the coder was uncertain what code to assign.

from the second (series of) incident(s).[18]

As mentioned in the introduction, the personal violence variable is the mix of three crimes of very different nature. Crime suffered by strangers, crime suffered by acquaintances and domestic crime. Thus, it may be properer to treat them as three separate crime categories. Fortunately, this information is also given in the Victim Forms data set and three separate dummies or count variables can be created.[19] This will be well discussed in Section 5.

Finally, for each incident of crime the information whether it is perceived as a racially motivated crime, together with the reason why it is perceived as such, is available. Therefore, this information can be used to control for 'perceived' racially motivated crime.

Although in the empirical analysis I focus on burglaries, personal thefts and violent crime, the distribution of all outcome variables is presented in Table 2. The full distribution of the violent crime variables is presented separately in Table 22. There are two main issues that deserve a brief discussion. Firstly, there is a very large concentration on the zero for most of the variables. Secondly, there are few cases of victims who reported extreme number of crimes. For instance, in variable Personal Theft there is only one person above ten crimes, who actually reported 97 crimes, or, for Inside Burglary there are eight people that reported between 70 and 100 crimes. In this table for ease of exposition we cap the crime count at ten plus more. Count data models are very sensitive to these cases, particularly when the positive counts are too few to identify the parameters assumed to affect the conditional mean. Someone would think of dropping these cases because they could be considered as highly unreliable. However, this practice would result in potential sample selection bias. Therefore, in Section 8 we also use several modified count data models that are both (in a sense) more robust under extreme cases and more appropriate to explain the observed distribution of victimization incidents. Finally, it is also clear that the dispersion of most variables is very high. Therefore, the Negative Binomial distribution that allows for over-dispersion may be more appropriate to fit the observed data.

Descriptive statistics of the outcome and explanatory variables are presented in Table 3. The means of all variables are separately given for native and immigrant groups in order to have a first indication on the victimization and socioeconomic differences between immigrants and natives. It must be noted that the immigration status variable is created as a dummy that takes the value 1 if the respondent or the HRP is not born in the UK. Moreover, the information of how many years the respondent lives in the UK can be exploited to examine assimilation patterns of the immigration-native victimization estimated differences. This will be examined in Section 6.

A first look at the raw data shows that there are victimization differences between immigrant and native groups, although they are very small in most cases. Regarding acquisitive crime,

[18]For instance, consider a case where a victim suffered 15 assaults without injury (1^{st} victimization form) and 5 assaults again without injury (2^{nd} victimization form). The difference between these two series of crimes may be that, for instance, the first series of assaults were committed by an acquaintance whereas the second series by a partner. Therefore, although these two crimes at the end take the same code (number 13), two different victimization forms are assigned. To construct the count of assault without injury for this individual we need to sum the count from the 1^{st} victimization form and the count from the 2^{nd} victimization form.

[19]The 'do' files (Stata® format) for the creation of dependent variables from the Victim Forms data set are available from the author upon request.

both household and personal, we can see that both the probability of victimization and the mean victimization are higher for immigrants, apart from Outside Burglary (and Outside Thefts or Other Thefts).[20] Moreover, Home Criminal Damage is slightly lower but Vehicle Criminal Damage is slightly higher for immigrants. Concerning Violent Crime, which is the crime group most discussed in this study, we can see that immigrants are less victimized. However, the picture is different if we break violence into the categories discussed before, as immigrants are much less victimized by acquaintances and family members, but slightly more by strangers.

Concerning the explanatory variables, first note that the means for the respondent's and the HRP's characteristic are given separately. This is because in personal crime analysis it is more appropriate to use the personal characteristics, but in household crime analysis it is more appropriate to use the household characteristics. The main observed differences between immigrants and natives is that immigrants are younger (which can be considered mainly as a measure of *exposure*) and that they are relatively more concentrated in London, urban and inner city areas; but most importantly that they reside in relatively more deprived areas (which can be thought of as *proximity* measures).[21] Moreover, immigrants are more married, more of nonwhite ethnic groups, less owners of their homes and they reside relatively more in flats (mainly *exposure* measures). They also live fewer years at their current home or area (which is a measure of *social guardianship*) and finally, they possess fewer cars (measure of *exposure*). There are no important differences in income and education.

It is important to note that, as the research question mainly concerns the immigration-native victimization link, in the next section our discussion will be concentrated on the effect of immigration dummy. Discussion of the effects of the other explanatory variables will not be given in the main text but as a note at the end of each subsection.

Finally, notice that for some of the independent variables there are many missing cases. Dropping all these cases would result in losing too much information. Therefore, a dummy is created for each variable that contains a considerable number of missing cases that takes the value one if the particular variable displays a missing value and zero otherwise. Thus, these dummies intend to absorb the effects of the group 'missing cases' of each characteristic on the dependent variables.

[20] Here we do not discuss statistical significance of the differences as these descriptive statistics are used just as a first indication.

[21] The **Deprivation Index** is the "Multiple Deprivation Index of England and Wales" for 2007, constructed as a weighted mixture of the individual deprivation indices (income deprivation, employment deprivation, health deprivation and disability, education, skills and training deprivation, barriers to housing and services, living environment deprivation, and crime deprivation index) provided by the Department of Communities and Local Governments for England and Welsh Assembly Government for Welsh. Very briefly, this index, that takes integer values from 1 to 10, provides a measure of multiple deprivation at the Lower Super Output Areas (LSOAs) level by considering some indicators of deprivation. These values indicate the decile of deprivation in which someone scores. For example, if someone scores at the 7th decile, only 30% of the population resides in more deprived areas. Each respondent, depending on the small level area that he/she resides, is matched by the Home Office with the corresponding decile of this variable. For more information on these indices refer to Noble at al (2008). In the empirical analysis I include this variable as an 1 - 10 integer index that measures the effect of scoring at a one decile higher on the probability of victimization.

4 Risk of Household Crime

In this section simple Probit results for household crime are presented. As discussed in the previous section household crime was separated in five mutually exclusive groups. However, here mainly the results of Inside Burglaries plus Attempts and Outside Burglaries plus Attempts are presented. The results of the other variables are briefly discussed in the second subsection. Full results are available from the author on request. The explanatory variables affecting the outcome variables are assumed to be the same for both crime groups.[22]

In the results that follow, four specifications of the conditional mean are presented. In specification 1 the effect of the HRP being an immigrant on the likelihood of victimization is considered without taking into account for socioeconomic differences between immigrants and natives. In specification 2 some important *proximity* measures are controlled for. In specification 3 some important characteristics of the HRP are also included, which are thought in literature to be associated mostly with the risk factor of *exposure*. Finally, in specification 4 some extra important household characteristics that are theoretically associated with *exposure*, *attractiveness* and *guardianship* are used.

4.1 Inside Burglary

Before discussing the results it is interesting to note that in 81% of Inside Burglaries plus Attempts the victim did not know the offender, whereas in only 10% of the cases the incident happened because of preexisted personal relationship/history between the victim and the offender. Therefore, although there are a few cases where interrelations and interaction between victim-offender matter, inside burglary can be considered to a high extent as "random" where criminals solely target the property without interest in the household composition and without intentions to victimize household members. Moreover, notice that in most cases, criminals' information about interior properties is limited, so that the value of the interior properties would not be a large factor for the risk of victimization. Instead, *attractiveness* is approximated by the external household characteristics.

According to the above, we would expect that if a relationship between immigration and inside burglaries exists, it is not because criminals prefer targeting immigrants' properties, but because immigrants' household characteristics are associated with more or less victimization, as discussed in Section 2. These characteristics refer to both direct household characteristics such as location and external condition, and indirect characteristics associated with the four risk factors, such as HRP's age, marital status, or how many hours the house is left unoccupied. Therefore, we would expect that a potential association would fade out if we were able to control for the characteristics that make immigrants' properties more or less likely to be attacked.

The Probit results are presented in Table 4. First of all, we can see that the likelihood of victimization increases if the HRP is an immigrant. The marginal effect is 0.74 percentage points

[22]Thus, we assume that the factors that affect the criminal opportunity structure through their effects on the four risk factors are generally the same for the two crime variables.

(which is statistically significant at 1% significance level) which is fairly large in magnitude if we bear in mind that the probability to suffer an inside burglary is 2.99% for immigrants and 2.25% for natives, a relative effect of 33%.[23] Note that the result is almost identical if we control for respondent's immigration status rather than HRP's immigration status. This was expected since it is highly possible that if the respondent is an immigrant the HRP is also an immigrant.[24]

Hence, dwellings in which the HRP is an immigrant are disproportionately victimized. However, a major part of this difference can be explained by the fact that immigrant disproportionately reside in urban areas where the deprivation index is much higher, two factors that are highly associated with the risk of inside burglary.[25] Moreover, from specification 3 it is clear that the rest of the difference is explained by HRP's basic characteristics indirectly associated with *exposure*, *attractiveness*, and *capable guardianship*. The association even becomes negative if we include the extra controls of the fourth specification.[26]

4.2 Outside Burglary

Outside Burglaries plus Attempts (burglaries in non-connected domestic garages/outhouses) are considered separately due to the following two reasons. Firstly, as immigrants disproportionately reside in flats or maisonettes, they probably possess fewer outside properties, such as non-connected to the main house garages, outhouses, storehouses and conservatories.[27] Therefore, controlling for other characteristics, the risk of outside burglary is expected to still be lower for immigrants. Unfortunately, information of outside properties is not given in the BCS. Secondly, outside properties can be considered by offenders as "safer" targets because of lower *physical* and *social guardianship*. Note that, using the Victim Forms data set, we can see that in 96% of the cases the criminal was a stranger and that in 99% of the cases the incident could not be attributed to previous personal history or relationship.[28] Hence, the same argument in favor of "randomness" used for Inside Burglaries holds here as well. Finally, notice that for this crime

[23] The standard errors of the marginal effects are calculated using the delta method (command 'nlcom' in Stata®).

[24] The tetrachoric correlation coefficient (see, Edwards and Edwards, 1984) is 0.9841.

[25] The marginal effect decreases to 0.29 percentage points, which is statistically insignificant.

[26] Regarding the effects of the variables from specification 4, we can see that if the HRP is older, married, employed and owner of home, the victimization risk falls. However, the gender of the HRP does not affect risk of victimization. For the rest of the coefficients in specification 4 we have the following relationships: as the perceived condition of the house increases, risk of victimization also increases. Moreover, condition of the dwelling relative to the other dwellings in the neighbourhood is important as both better and worse condition houses are of higher risk of victimization. Moreover, detached houses, and properties located on main or the side of the road are associated with more crime. Number of adults in the house and hours that the house is left unoccupied have no effect. On the other hand, if the respondent is a lone parent the risk of victimization increases. The longer the respondent resides in the same house the lower the likelihood of an inside burglary. In addition, if the property is in a neighborhood Watching Program the risk of victimization decreases (significant at 10%). The joint effect of income dummies, having less than 10,000 pounds of annual income as the reference group, is significant at 1% with 50+ group being the only group associated with more crime than the base group (significant at 10%). Finally, education dummies are jointly significant at 10%, with more crime for the ones with higher education.

[27] We need to stress that theft of outside properties and car thefts are not included in outside burglaries but they are treated separately.

[28] Of course, this might be because in most of outside burglaries it is highly likely that the victim had no contact with the offender, and therefore, could not be able to evaluate whether he/she knew the offender.

category we observe very few positives (99% of zeroes).

The results are depicted in Table 5. Despite the very high proportion of zeroes, Table 5 shows that the estimated effect of the HRP being an immigrant is very robust across all specifications and statistically significant at 5%. To evaluate the magnitude of this difference marginal effects are calculated for a 'representative' household. For example, in specification 2, evaluated for a household that is located in an average deprived area, in the inner city of an urban area in London, the probability of an OUTSIDE BURGLARY PLUS ATTEMPT is around 0.3 percentage points lower for households in which the HRP is an immigrant, a relative effect of around 60%.

Thus, even though immigrants live in relatively more deprived areas, they face a much lower probability of victimization. This may be attributed, as mentioned before, to the fact that immigrants possess fewer domestic outside properties. Unfortunately, there is no information on non-connected domestic outside properties and therefore, we are not able to test the above argument (even though we control for the fact that immigrants are more likely to live in flats). However, a zero-inflation (ZI) count data model could be relevant in this case (see, Mullahy, 1986, and Lambert, 1992). According to the ZI model some households will never experience an outside burglary just because they do not own any outside properties. It is interesting that, in accordance with this previous argument, ZI models for counts show that the immigration status coefficient is positive in the zero-inflation equation and significant at least at 10% significance level in most specifications. A zero-inflated Probit model was also employed, whose log-likelihood function resembles the log-likelihood of the MisProbit model presented in Papadopoulos (2013) if the one inflation probability is constraint to be equal to 0. Although the behaviour of this model in terms of estimation was not trustworthy, its results also indicate that the proportion of immigrants in the zero inflation category is more than the proportion of natives and significant at 5% level of significance. All results are available from the author on request.

Moreover, we could think that earlier immigrants are better settled and therefore, their outside properties would be more similar to natives' ones. Thus, we expect to observe a lower risk of outside burglaries for earlier immigrant with an assimilation pattern as the number of years in the country increases. Unfortunately, NUMBER OF YEARS IN THE COUNTRY is not provided for the HRP, but we could approximate it with respondent's NUMBER OF YEARS IN THE COUNTRY since, as noted in the previous section, the results were very similar when using the variable IMMIGRANT instead of HRP IMMIGRANT. The results, which are presented in the first two rows of the lower part of Table 5, are quite supportive of the above argument.[29] We can see that when we include a linear trend for the number of years of an immigrant in the host country, more recent immigrants are associated with a much lower probability of victimization (even lower than before) and that this probability converges to natives' one as years in the country increase (although the marginal effects show that it takes more than 40 years for immigrants to assimilate to natives' probability of outside burglary victimization). Note that the "assimilation" coefficient is insignificant in specifications 1 and 2 because we do not control for age, as immigrants that are more years in the country are relatively older, and older people are associated with lower

[29]Here, only the coefficients of interest are presented. Full results are available upon request.

victimization risks. Once we control for age, the coefficient of immigration dummy increases in magnitude and the "assimilation" coefficient becomes significant at 5% significance level.

Finally, note that most of the other regressors have an insignificant effect on the probability to suffer an outside burglary.[30]

4.3 Remaining Household Crime Types

In this subsection the main results of the association between immigration and the risk of victimization for Vehicle Thefts, Household Thefts and Criminal Damage are briefly discussed. The results are not presented but are available from the author on request.

To begin with, contrary to burglaries, Vehicle Thefts are much more often, as the probability of victimization in the raw data is 6.53%. Therefore, the estimates obtained for this crime group are much more precise. Once more, we expect that holding everything else constant, immigrants would experience a lower risk of vehicle thefts just because they own fewer vehicles. However, as opposed to outside burglary, in this case we have information on both the number of cars a household owns and on ownership of motorbikes and bicycles. The results show that immigrants face a higher risk of vehicle thefts (statistically significant at 1%) even though they own fewer vehicles, if we do not control for demographic disadvantages of immigrants. Thus, the coefficient of the effect of immigration status on vehicle crime increases once we control for this fact by including the natural logarithm of vehicles as a regressor and considering only the population that possesses vehicles.[31] However, as expected, if basic demographic differences between immigrants and natives are controlled for, the difference in the likelihood of victimization dissapears.

Household Thefts consists of Inside Thefts (0.25% positives), Outside Thefts (2.62% positives), Other Household Thefts that do not fall within these two categories (1.76% pos-

[30]For the coefficients in specification 4 we have the following relationships: only relative condition affects victimization, as the better the condition relatively to other houses, the higher the risk of victimization. There is no effect for worse condition. The dummies for the type of the house have no joint effect. Being located in a main road increases the risk of victimization but being in a side road does not affect it. Number of Adults has no effect either. On the contrary, as for inside burglary, lone parents' households experience a higher risk. Moreover, there is no effect for, Hours Unoccupied, Years at Home and Years in Area, neighborhood Watching Program and income dummies. Finally, education is jointly significant at 1%, with more crime for more educated people (more than a-levels).

[31]The reason why we include the number of vehicles in the natural logarithm form is the following: firstly, it is important to note that any binary choice model could be thought of as a censored at 1 crime count data model. For example, in the Poisson case, the probability of the zero outcome is $e^{-\lambda}$ and the probability of a positive is $1 - e^{-\lambda}$ where λ is the Poisson conditional mean. Thus, the structure of the conditional mean of the binary model should be consistent with the structure of the conditional mean of a count data model. As it is very common in count data models, in order to ensure nonnegativity we consider the mean to be given by $\lambda_i = \exp(\mathbf{x}_i'\boldsymbol{\beta})$. Moreover, it is natural to assume that the risk of suffering a vehicle crime is proportional to the number of vehicles someone possesses (in the same way we model cases where different individuals are exposed on the outcome y for a different time interval), since the number of vehicles can be considered as a direct measure of *exposure*. Thus, if N is the number of vehicles someone possesses, the mean in this particular case is given by $\frac{\lambda_i}{N} = \exp(\mathbf{x}_i'\boldsymbol{\beta}) \Rightarrow \lambda_i = N \cdot \exp(\mathbf{x}_i'\boldsymbol{\beta})$. Therefore, the number of vehicles should be included in the regression framework as the ln of N, so that $\lambda_i = \exp(\mathbf{x}_i'\boldsymbol{\beta} + \ln N)$. From the last expression it is clear that we cannot include the households with zero vehicles. Intuitively, considering only the population that possesses vehicles, we directly control for the zero-inflation probability which is the probability of not suffering vehicle crimes just because of no possession of any vehicles (No *exposure*).

itives) and Attempted Thefts (0.16% positives).[32] The results indicate that immigrants do not experience a higher risk of being victims of Household Thefts even thought they have some demographic disadvantages (the coefficient is 0.002 and very insignificant). Therefore, when we control for demographic differences, the coefficient becomes negative and significant at 5%. It has to be stressed that these results are driven by Outside Thefts, as it is the variable with the highest proportion of ones. If we break household thefts in the three categories we observe the following: for Inside Thefts immigrant coefficient is always positive but insignificant in all specifications; for Outside Thefts it is negative but insignificant in specifications similar to 1 and 2 of Tables 4 and 5, but negative and significant at 10% if we include further controls; finally, for Other Thefts it is positive and insignificant in specification 1, but negative and insignificant in specifications 2, 3 and 4. Thus, immigrants face a lower probability of Household Thefts probably because they do not own many outside properties, or because they are more capable of protecting and monitoring their outside properties.

Finally, the nature of Criminal Damage is very different, since vandalism is an expressive crime, as opposed to the other crimes discussed above (acquisitive crimes). Criminal Damage includes Home Criminal Damage (2.48% positives), Vehicle Criminal Damage (5.37% positives), Other Criminal Damage (0.11% positives) and Arson (0.001% positives). The empirical analysis shows that, as for Household Thefts, although immigrants stay in disadvantageous areas, they experience the same risk of vandalism. Therefore, the coefficient of immigration status becomes negative and significant at 5% in specifications 3 and 4 (but not significant in specification 2). Further analysis shows that the previous effect is driven by the effect of immigration on Home Criminal Damage, as there is no relationship for Vehicle Criminal Damage (as it was the case for Vehicle Theft).

5 Risk of Personal Victimization

In this section the results of personal victimization are presented. First of all, there is one essential factor that is different between personal and household victimization; personal victimization entails personal contact between offender the victim. As a result, personal characteristics of the victim might directly affect the criminal action. The implications of this crucial difference on the immigration-victimization relationship can be quite important. This is mostly because, as potential offenders directly observe potential victims, they are able to approximately determine the ethnic background of the potential victim. Thus, the fact that someone is an immigrant might have an effect on the victimization probability even after controlling for a large set of observed individual characteristics, if there are still immigrants' characteristics associated with personal victimization that are observed by potential offenders but unobservered in the data.

[32]The differences between a burglary and a household theft are explained in detail in Bolling, Grant, and Donovan (2008). Very briefly, inside thefts consist of the cases where there was a theft by a person who was in the house with the consent of households members. Outside thefts consists of thefts of properties outside the house without any sign of outside burglary. Other household thefts include all other categories of household thefts excluding personal thefts.

For instance, immigrants may appear as more vulnerable and therefore, they could be considered as an easier and 'safer' target.

In addition, there is also a crucial difference between the two main personal crime types, Personal Theft and Personal Violence, which indicates that they should be treated separately. Personal theft is an instrumental type of crime whereas violent crime is an expressive type. Therefore, contrary to personal theft, as discussed in Section 2, a violent action in most cases requires personal interaction between the potential victim and the potential offender. This should not be translated as prior history in the victim-offender relationship, as there can still be interactions that generate a violent act even for individuals that were unknown to each other prior to the incident, such as brawls or arguments in pubs and bars. According to this, there might even be cases where the victim is at the same time an offender, which is unlikely for personal theft. Therefore, unobserved personal behavior is much more important for violence than for personal theft.[33] The potential offender observes the potential victim and once a set of information is obtained, an evaluation of the expected utility follows. If the expected gains are higher that the expected costs the individual commits the crime.[34] In the first subsection the risk of personal theft is examined, whereas the analysis for violent crime follows in the second subsection.

5.1 Risk of Personal Theft

First of all, it is important to note that in the present study robberies are considered as personal thefts although they entail violence. I examine robbery in this category rather than in violent crimes because primary target of the offender is to acquire victim's valuables and not just to hurt the victim. Personal Thefts (1.59% positives) consists of Robberies plus Attempts (0.42% positives), Snatch Thefts from the Person (0.15% positives), Other Thefts from the Person (0.73% positives), and Other Attempted Personal Thefts (1.39% positives).[35]

Table 6 presents the results in four specifications. Predicted probabilities and marginal effects are also presented. From the first specification, where the effect of being an immigrant on the risk of a personal theft is examined without controlling for other characteristics, we see that the probability of victimization is much higher for immigrants (61.2% higher). As shown in specification 2, this difference cannot be totally explained by immigrant-native differences in some important demographic characteristics (the relative effect of 34.9% is still very high).[36]

[33]Note that 94% of personal thefts were committed by strangers, 98.8% thefts did not happen because of prior history/relationship between the offender and the victim, and from the few cases where the victim consider himself/herself as responsible for the action (only 6% of the incidents) there is no incident where the victim provoked the offender.

[34]For a formal model on the decision to commit property crimes see, Papadopoulos (2013).

[35]You can notice that adding up the 5 personal theft groups together you obtain a probability of victimization which is higher than the probability to suffer a personal theft as a whole. This is because it is possible that a person suffers more than one type of personal theft.

[36]The marginal effects are calculated for a white male, between the age of 36 and 45, who stays in an urban area where the deprivation index takes the average value. It seems that 'ethic group' matters for personal theft.

However, the third specification reveals that immigrants are more likely to become victims of personal theft because they disproportionately reside in London, which is, according to the estimates, the place with the highest risk of personal theft.[37]

5.2 Risk of Violence

Violent Crime (2.54% positives) includes Assaults with Serious Wounding (0.21% positives), Assaults with Other Wounding (0.55% positives), Common Assaults (1.6% positives), and Attempted Assaults (0.32% positives). We need to stress that violent crimes with sexual motive and robberies are not included in this group. As discussed before, violence is an expressive type of crime where interrelations and interactions between potential victims and potential offenders are vital. As an indicator of this, the Victim Forms data shows that in 23.03% of the victimization incidents the victim knew the offender casually, and in 34.44% he/she knew the offender very well. Moreover, in 27.34% of the cases the incident happened because of previous history/relationship between the victim and the offender. Finally, in 6.31% of the cases (81 incidents) the respondent considers himself/herself as being responsible for the action, while in the 65.43% of these 81 incidents there was provocation by the victim, which means that probably the victim initiated the action. Therefore, unobserved (in the data) characteristics associated with *routine activities* and *lifestyle-exposure* could be important on explaining remaining differentials in the immigration-victimization relationship.

The results for this crime category are presented in Table 7. Specification 1 shows that, without controls, immigrants face a lower risk of victimization but the difference is statistically insignificant. However, the marginal effect is significant at 10%. As it is clear from specification 2, victimization decreases considerably with age, and since immigrants are relatively younger, controlling for age results in increasing the magnitude of the immigration status coefficient. The marginal effect, which is significant at 1% level, shows that being an immigrant decreases the probability of a violent incident from 3.61% to 2.48%, a difference of 1.13 percentage points. According to the estimates of specification 3, controlling for geographical factors does not alter the immigration-violent victimization estimated link. Note that the estimated coefficient increases even further in magnitude if we do not include regional dummies. This is very interesting because London is the place whose residents experience the lowest risk of violent victimization, as opposed to personal theft, where London's residents faced the highest risk of victimization. Finally, it is quite important that the effect of immigration preserves its magnitude and significance even

Black individuals experience a higher risk, while Asians, Chinese and Others experience a lower risk.

[37]For specification 4, the effects of the variables whose estimates are not presented in the table are the following: education dummies are jointly significant at 1% (having no qualification as the baseline group), with more than a-levels people being the most victimized group. Income dummies are jointly significant as well, but the relationship is not very clear. People of the lowest income category (10,000 or less) face higher risk than the 10,000-20,000 income category. The group from 20,000-40,000 face lower risk but the effect is insignificant, while the group 40,000-50,000 experience more risk but the effect is again insignificant. Finally, the group more than 50,000 experience higher risk but still insignificant. For the dummies of employment status (where employed people is the baseline dummy) and marital status (with married people being the baseline dummy), employed and married people face the lowest victimization risk. Finally, owners experience a lower risk than renters.

when we use some other observed characteristics associated with risks of violence.[38]

Furthermore, in Table 8 we present the results of the same specifications once we include dummies for ethnic background. As expected, inclusion of ethnic dummies affects the immigration status coefficient (which becomes more significant in specification 1, but less significant in specifications 2-4), since immigrants are disproportionately from ethnic minority groups. This can be also seen by the marginal effects.[39] Concerning the effect of the ethnic dummies, although it seems that Asian individuals, and to a smaller extent Black ones, experience a lower risk of victimization relative to White individuals, their joint effect is insignificant. Even in the last specification where both the effect of Asian and Black individuals relative to White ones is significant at 5%, the joint effect of the ethnic group dummies is not statistically significant (the p-value from the Wald test is only 0.123).

Therefore, it seems that immigrants experience a lower risk of violent victimization due to immigrant-specific unobserved characteristics. A general explanation for this could be that immigrants set strategies that correspond to unobserved differences in *routine activities* or *lifestyle-exposure* associated with lower criminal activity. For instance, immigrants may avoid socializing in places where there is a high risk of violence, such as pubs or clubs,[40] or that they exhibit a less provocative behavior.[41] This result is consistent with the results of Papadopoulos (2013) who shows that immigrants are slightly less likely than natives to commit violent crimes. Thus, immigrants in general follow some crime-avoiding lifestyles which result to a lower *exposure* to violent crime activities than natives and as a result, they face a lower probability of both committing and suffering a violent crime. Moreover, a part of the estimated difference could be explained by the following hypothesis, also consistent with the results of Papadopoulos (2013) and closely

[38] The effect for the rest of the controls in specification 4 is the following: the education dummies (where baseline group is no qualification) are not jointly significant. However, it seems that the risk of victimization increases with higher education. Being married lowers the risk while being single has the highest risk. Unemployed individuals have higher risk than employed ones, while inactive individuals endure the same risk. Regarding income dummies effects (where the base is less than 10,000 pounds), all groups suffer lower violence than the poorest group, however, the statistical significance decreases as income increases. In addition, the risk increases for lone parents and bigger households. Also note that the marginal effects are evaluated for the following representative individual: a male, between 35-44 years old, residing in an average deprived urban area in the East of England, who has a-levels qualifications, and also he is married, employed, owns the place he lives and finally belongs to a family with 2 household members. Finally, we need to mention that there are two variables derived from the questions, "how often have you visited a pub in the last month" and "how often have you visited a club in the last month", which are asked by the conductors to be used as a proxy for *exposure*. However, this information can be considered as a poor measure of *exposure* if we are not able to control for day-life activities and other activities associated with more or less *exposure*. Thus, this regressor is measured with error for representing a *lifestyle-exposure*, which attenuates the immigration coefficient since there is a strong and statistically significant negative association between being an immigrants and going to pubs and clubs (being an immigrant decreases the probability of going to clubs or bars by around 18 percentage points, a relative effect of around 53%). Nevertheless, the coefficient of immigration status is still significant at 5% in specifications 2, 3 and 4.

[39] The marginal effects are calculated for the same individual as before, plus the extra characteristic that he is white.

[40] According to the BCS data 35% of all immigrants, but 53% of all natives, have been to a pub or a bar during the month prior to the interview.

[41] According to the BCS Victim Forms, from the 980 victimization incidents where the victim finds himself/herself as responsible for the incident, we observe that only 6.32% of immigrant victims provoked the offender, but 8.93% of native victims provoked the offender. Note that here I include all types of crime. If we consider violent crime only, these figures change to 50% for immigrants and 65.79% for natives, but note that there are only 4 violent crime incidents where an immigrant considered himself/herself as responsible for the incident.

related to the one above. If we accept that immigrants socialize mostly with other immigrants, and if we also assume that immigrants socialize with the same number of people as natives do, the probability of violent victimization would be lower for immigrants just because immigrants are less violent.

However, as discussed in Section 3 violent crime is composed of three different types: Domestic Crime, Crime by Acquaintances, and Crime by Strangers. Therefore, it is important to investigate the immigration-victimization link for each violent crime type separately. This analysis follows in the next 3 subsections.

5.2.1 Domestic Crime

In the present study, Domestic Crime refers to violence experiences within a family. Note that the proportion of positives is really low, as only 0.51% of the respondents report that they have experienced domestic violence.

The Probit results are presented in Table 9 in four specifications.[42] The coefficient of the marital status dummies are also presented as they seem very important in explaining domestic crime. We can see that the likelihood of an immigrant being a victim of domestic violence is much lower in all specifications. Being an immigrant almost halves the probability of domestic violence.[43] Someone would argue that this is driven by the fact that some immigrants, particularly younger or more recent ones, leave their families back as they intend to work for a few years and return back. It might also be the case that due to cultural differences immigrants might be less willing than natives to report domestic violence.[44] Both issues will be examined in the sensitivity analysis section.

From Table 9 we can also see that men are less victimized than women as expected. In addition, it is noteworthy that divorced and separated individuals face the highest risk of victimization. Thus, women get victimized by ex partners during the 12 months prior to the interview, or victimized individuals tend to move forward incidents that happened long time ago, or married people for some reasons tend to under-report disproportionately. Finally, it is worth mentioning that the deprivation index is not associated with higher crime once we control for marital status.[45]

[42] Ethnic dummies are not used for domestic crime, as they do not affect the probability of domestic crime even when we do not control for immigration.

[43] The marginal effects are evaluated for a female, between 36-45 years old, with all other characteristics the same as in the previous subsection.

[44] If immigrant families are in a sense more "traditional" or more patriarchal, fear of reprisal could be higher for them, resulting in higher under-reporting.

[45] With regard to the effects of the other variables we have the following relationships: education dummies have no joint effect. Income dummies are jointly significant with poorest people being the group associated with the highest risk of victimization. Lone parent has a positive and significant effect even after controlling for marital status and number of household members. However, bigger households are not associated with higher or lower victimizations. The effect of regional dummies is significant at 5%, London being the region with the lowest risk of domestic victimization.

5.2.2 Crime by Acquaintances

Crime by Acquaintances refers to crime suffered by people who are familiar to the victim, but not family members. Only one percent of respondents suffered a crime by familiar people. As for domestic crime, prior history is also important for this type of crime. As an indication, in around 30% of crime by acquaintances prior history was responsible for the incidence and in 55% out of the 36 cases where victims consider themselves as responsible for the incident,[46] the victim provoked the offender.

The results, depicted in Table 10, are striking. From specification 2 we can see that natives are more than 100 percent more likely to suffer a crime by acquaintances, once we control for basic demographics.[47] The immigration status coefficient preserves its significance and magnitude even under a rich set of observed characteristics. In specification 4, where we also include controls for ethnic status (as now ethnicity dummies have a joint significant at 5% effect), immigration coefficient loses some of its significance and magnitude (as now being an immigrant decreases the probability of victimization by around 60%) as anticipated, but it is still significant at 10%, which is still important given the very few zeroes in the dependent variable (even though the data set is quite large).

This result is consistent with the findings of Papadopoulos (2013). If we accept that acquaintances of one ethnic group consist in a high proportion of people from the same ethnic group, we expect a high proportion of immigrants' (natives') acquaintances to be immigrants (natives) as well. Since immigrants are less prone to violent crime as offenders, we expect that immigrants would be less likely to suffer crimes by acquaintances relatively to natives. Moreover, if immigrants a crime-avoiding behavior, not only would they not initiate arguments or fights, but they would also avoid socializing in “dangerous" places or with “dangerous" people. On the other hand, it could also be that immigrants are less likely to suffer a crime by acquaintances just because they have smaller networks of acquaintances (“network effects”), a feature directly associated with *exposure*. This hypothesis will be examined in the next section. Finally, as for domestic crime, we cannot exclude the possibility that immigrants might be less willing than natives to report crimes that suffered by friends or other familiar individuals.[48]

5.2.3 Crime by Strangers

Crime by Strangers involves brawls in pubs and bars (31% of the cases), arguments and fights on the streets or in public means of transportation and so forth. In the data, 1.09% of respondents experienced a victimization incident by a stranger. Although this crime can be considered as

[46]The victim believed that he/she is responsible for the incident in 36 out of 507 cases (7.1%).

[47]The marginal effects are calculated for a person between 36-45 years old, and rest of characteristics the same as the individual in Violent Crime results.

[48]Note that the effects of the variables whose coefficients are not presented in Table 10 are as follows: regional dummies affect victimization significantly, London being the place with the lowest victimization. Risk also increases for bigger households. The effect of income is significant as well, and the risk of victimization becomes smaller as income increases. On the other hand, education is jointly insignificant. Finally, married people and home owners face a lower risk of victimization by acquaintances.

more "random" than crime by acquaintances and domestic crime, interactions between offenders and victims are still important. For instance, it is not very likely that someone will be attacked while walking down a street without any reason, unless the primary target is to acquire victim's property which is, however, recorded as a robbery. According to our data, only in 17 out of 529 incidents the victim considered himself/herself as responsible for the action (2.26%), 9 of which the victim provoked the offender (52.94%).[49]

The results for this crime category are presented in Table 11. Contrary to the other two types of crime, immigrants are equally likely to suffer a crime by a stranger, even after controlling for disadvantageous characteristics of immigrants. Thus, the results of TOTAL VIOLENCE were driven by domestic crime and crime by acquaintances. This is in contrast with the "crime-avoiding" social behavior of immigrants discussed in the previous subsection, since we would expect to observe a similar pattern between being an immigrant and crime by strangers, and being an immigrant and crime by acquaintances, if immigrants do avoid criminal actions in general.

Thus, this finding raises some important questions. Why do we observe a significant negative immigrant-victimization association for DOMESTIC CRIME and CRIME BY ACQUAINTANCES, but no association for CRIME BY STRANGERS? Is it that immigrants actually **do not follow** the "crime-avoiding" social lifestyles but we still observe a negative effect for crime by acquaintances and domestic crime because: 1) immigrants are less willing than natives to report domestic violence and crime committed by acquaintances; 2) immigrant have fewer household members or smaller networks of acquaintances compared to natives. Or is it that immigrants actually **follow** the "crime-avoiding" social lifestyles but we do not observe a difference for crime committed by strangers because immigrants are more likely to suffer more racially motivated crimes (RMC) compared natives, a crime that is usually committed by strangers and it is traditionally associated with ethnic minorities? These issues are examined in the next section.

6 Sensitivity Analysis

Before presenting the results of this section, we need to stress that henceforth we will be controlling only for the following basic demographic characteristics: GENDER, AGE, DEPRIVATION INDEX, REGIONS, URBAN and INNER CITY. Thus, all the following results look at the differences in the likelihood of victimization between natives and immigrants, if these two groups exhibited the same basic demographic characteristics.

6.1 Controlling for Correlated Errors

In the previous section we treated the three violent crime categories as being independent from each other. However, it may be properer to take into account the fact that someone who suffered

[49] Note also, that 41.9% of the crimes by strangers happened because the offender was under the influence of alcohol or drugs and 18.5% because of an attack by young people, teenagers or mindless vandalism.

a violent crime of one group may also be likely to suffer a crime a violent crime of another group. Thus, conditional on a set of the observed characteristics, there might be some common individual unobservable factors across the three crime groups. Accordingly, we could use a model that allows for correlated errors across the three crime equations, similar to the Seemingly Unrelated Regression framework (see, Parks, 1967). This can be done by using a Trivariate Probit model which might result in efficiency gains as it exploits the information that some sets of unobserved characteristics appear in all equations (see, Greene, 2011, for a formal representation of Bivariate and Multivariate Probit models).

A complexity here is that, although there are algorithms to evaluate univariate and bivariate standard normal integrals, these algorithms cannot evaluate M-variate normal integrals (see, Greene, 2011). On this direction, a simulation-based integration has been developed (see, Cappellari and Jenkins, 2003). Therefore, for the purposes of this analysis a simulated maximum likelihood three-equation Probit estimator that uses the **Geweke-Hajivassiliou-Keane** smooth recursive simulator is used (see, Terracol, 2002).[50] Obtaining estimates by using this estimator is time demanding and therefore, the number of draws to select is quite important. According to Cappellari and Jenkins (2003) this estimator is consistent when the number of draws and the sample size go to infinite. However, they find that a number of draws close to the square root of the sample size is a reasonable number to use. In my case, it is found that the estimated coefficients change very little if the number of draws is larger than 200.[51] The results of this model, which are presented in Table 12, are obtained using 300 draws.

We can see that the estimates of this model, both for the immigration coefficient and the coefficients of the other regressors, are very similar to the estimates when we treated the three crime group as independent but more precisely estimated. The only change is that the estimated coefficient of immigration status in the domestic equation loses a little of its magnitude. However, since this coefficient is more precisely estimated, its statistical significance remains the same.

It is also very interesting that we estimate a significant at 1% level positive correlation of the errors between Domestic Crime and Crime by Acquaintances equations, and between Crime by Acquaintances and Crime by Strangers equations, but no correlation between Domestic Crime and Crime by Strangers. This implies that, due to unobserved factors, individuals who suffer a domestic crime are likely to suffer a crime by an acquaintance as well than a person who does not suffer a domestic crime, and *vice versa*. The same holds for the pair, Crime by Acquaintances and Crime by Strangers, but interestingly, it does not hold for the pair Crime by Strangers and Domestic Crime. Moreover, we can see that a likelihood ratio test rejects the hypothesis that the three equations are independent.

However, as the estimated coefficients between the single equation Probits and trivariate Probit are very similar, and given that this estimation procedure is highly time consuming,

[50]To obtain these estimates the 'tribrobit' command in econometrics software Stata®, written by Antoine Terracol (2002) was used. A similar Stata® command that is generalized to account for a larger number of equations is written by Cappellari and Jenkins (2003).

[51]Only changes after the second decimal points of the estimates were observed. However, the estimated correlations among the error terms seem more sensitive to the number of draws selected.

we keep presenting the results of the conventional Probit models. Alternatively, the estimated correlations of the errors suggest that (the much simpler in terms of time and numerical intensity) bivariate Probit models between the two crime pairs could be used. However, even these results are very close to the ones obtained by conventional Probit models.[52]

6.2 Examining Differences in Reporting Behaviour

As discussed above, a reason why the effect of immigration status on crime suffered by strangers is different from the effect on crime by familiar people might be that immigrants under-report by more than natives crime experiences by familiar people. Thus, the question is: is it that immigrants do not hesitate to report crimes suffered by strangers (and thus, observing no differences in the risk of victimization) but hesitate to report crimes by acquaintances and family members? In the next two subsections, following two different strategies, we show that immigrants do not under-report, at least by more than natives. Firstly, we use self-reports on domestic violence and secondly, we exploit the available information on whether the partner was present during the face-to-face interviews. Both of them will provide important insights on differences between immigrants and natives reporting betavior.[53]

6.2.1 Use of Self-Completions on Domestic Violence

As mentioned in the introduction there is evidence that respondents under-report domestic crime (see, for example, Walby and Allen, 2004). Self-completions, as opposed to face-to-face interviews, are used as a technique to elicit more reliable responses to sensitive questions (see, Turner at al, 1998). For this purpose, people from 16 to 59 years of age were asked to self-complete a computer-based questionnaire for domestic crime. Therefore, a dummy SELF-COMPLETED DOMESTIC CRIME was constructed which takes the value one if the individual revealed (in the computer-based questionnaire) that he/she suffered a crime by any family member and zero otherwise. This variable consists of assaults and serious threats. Note that sexual abuse is not used. Regarding under-reporting the results are striking. Only 0.51% of the respondents reported a domestic crime in face-to-face, but 3.64% in self-completion interviews.[54]

Given that under-reporting is much lower in self-completions for both immigrants and natives, if in face-to-face interviews immigrants under-report by more than natives, we would expect that the effect of immigration status on SELF-COMPLETED DOMESTIC CRIME to be quite smaller

[52]These results are available from the author upon request.

[53]A third approach that uses two parametric models which are more appropriate under the presence of under-reporting was also followed for both binary and count data models. The binary model, which is based on Hausman, Abrevaya and Scott-Morton (1998), is the model presented in Papadopoulos (2013) under the name of MisProbit apart from the difference that the probability of over-reporting in the present study is set to zero. References for the count data models include Papadopoulos (2013) and Papadopoulos and Santos Silva (2008). The results of these models show that if anything, immigrants under-report by less than natives. However, these results were not very reliable, probably because of both the low proportion of positives and the noisy nature of victimization data. Thus, they are not presented in this study but they are available upon request.

[54]However, we must be cautious with this difference as the questionnaires between these two different types of interviews and the whole procedure followed to construct the two data sets are quite different (for details refer to, Bolling, Grant, and Donovan, 2008)

than its effect on FACE-TO-FACE DOMESTIC CRIME, as immigrants would now report more freely.

There is a small complication that does not allow us to directly use conventional Probit models though. This is because some individuals chose not to participate in the self-completion procedure. Controlling for a set of observed characteristics, are the individuals who are more likely to suffer domestic violence also more reluctant to participate (*selection on unobservables*)? Or, controlling for the same set of observed characteristics, does the likelihood of domestic victimization play no role on the probability to participate (*selection on observables*)?[55] If there is selection *on unobservables*, **Heckman-type Sample Selection Probit** models (Heckman, 1979) need to be used. Otherwise, the **Inverse Probability Weighted Estimator** (IPW), or even a simple Probit model under some circumstances, should be used.[56]

First of all, comparing immigrants and natives' participation rates, we find that immigrants' probability to participate in the self-completion procedure is much lower than natives' one. Only 5.58% of natives between 16-59 years old did not participate, but 13.98% of immigrants. In addition, there is an extra complication. Some people who accepted participation, for some reasons asked for the help of the interviewer to complete this questionnaire. These people did not answer the crime questions of the self-completed questionnaire. From people that accepted participation, 13.06% of natives did not complete the relevant crime questions while the 21.46% of immigrants did not complete them. Thus, altogether, 32.44% of immigrants did not complete the self-questionnaire compared to 17.91% of natives. Therefore, if people that did not participate are more likely to have been victimized than participants, and given that respondents report more truthfully in self-reports, then the coefficient measuring the immigrant-native SELF-COMPLETED DOMESTIC CRIME link would be downward biased.

In this subsection I use the estimator proposed by Van de Ven and Van Praag (1981), a maximum likelihood modified Probit estimator that is consistent and asymptotically efficient under sample selection is on unobservables. Two different specifications for the Sample Selection Probit are considered. In the first one, we treat people that accepted participation but did not answer the crime questions, as if their behavior towards the selection process is the same as the behavior of those who did not participate at all, and we use a sample selection model including in crime equation only people who self-completed the crime questions. In the second one, we exclude people that initially rejected participation and we keep only the sample of people who accepted participation. So, in the second model the selection process includes only individuals who accepted participation, whereas in the first case it includes all individuals between 16-59 years old.

[55] Formally, if s is the selection dummy taking value one if the individual participates and zero otherwise, $\mathbf{z}$ is a vector of observed characteristics which includes the set of explanatory variables that belong to the victimization equation $\mathbf{x}$ and possibly other variables that although they affect victimization we choose not to include them, and u is the error term in the victimization equation, the selection is on unobservables if $P(s_i = 1|\mathbf{z}, u) \neq P(s = 1|\mathbf{z})$ and on observables if $P(s_i = 1|\mathbf{z}, u) = P(s = 1|\mathbf{z})$. See, Wooldridge, 2007, for details.

[56] Actually, for the simple Probit estimator to be consistent we need that $P(s_i = 1|\mathbf{x}, u) = P(s = 1|\mathbf{x})$, so that the selection is on the explanatory variables only.

The results are depicted in Table 13.[57] As can be seen from this table four separate specifications are used. In the first specification we present the simple Probit estimates of face-to-face interviews for all respondents between 16-59 years old for the sake of comparisons. In the second specification a model that does not correct for sample selection for the sample of the individuals that contributed to the self-completions only is given. Finally, in specifications 3 and 4 we present the results of the two Sample Selection Probit models discussed above. First of all, we note that selecting our sample to individuals between 16 and 59 years only old does not seem to bias our estimates, since the estimates of specification 1 are similar to the ones obtained using the full sample. This will be the reference model for comparisons. Note also that the sample in specification 4 is different from the sample in specification 1, even though both models include all respondents between 16 and 59 years old. This is because there are some people whose answers on self-reported crime questions were recorded, for unspecified reasons, as missing cases.

It is well known that the sample selection models are better behaved if at least one exclusion restriction is imposed on the crime equation. Otherwise there is severe multicollinearity and identification is obtained only due to nonlinearity of the functional form. For this reasons, in model 4 we assume that OTHER PRESENT, a dummy variable that takes value one if someone else was present during the face-to-face interview and zero otherwise, and NO QUALIFICATION, a dummy that takes value one if no educational qualification is obtained and zero otherwise, are two variables that have an effect on the selection process but not on the crime process, conditional on the observable characteristics. The presence of someone else during the interview might affect the selection process as the respondent may feel some kind of pressure from the other members. For instance, in an extreme case, the partner may have prohibited the respondent from completing the self-report questionnaire. In another direction, the presence of others might indicate that respondents needed help during the interviews and therefore, did not answer the relevant crime questions. NO QUALIFICATION may affect participation, for example because less educated individuals face more difficulties in using the computer.[58] In specification 3 the dummy variable LANGUAGE DIFFICULTIES is used. Once respondents accepted self-completion, they replied to the question whether they have language difficulties, which is a major factor for asking help to complete the questions but not a factor directly affecting the likelihood of suffering a domestic crime. Note that this variable cannot be used in model 4, as answers on this question are conditional on accepting participation.

Table 13 gives some very interesting findings. First of all, we can see that immigrants are significantly less likely to participate and subsequently answer the crime questions than natives. Moreover, it is also clear that the variables used only in the selection equation have strong negative effects on the likelihood of selection. However, we notice that the immigration status

[57]According to the results of the previous section, marital status dummies are very important factors of domestic violence. However, the results of the models that include these dummies are very similar to the ones presented in Table 13 and therefore are not presented here (but are available on request).

[58]Note that the NO QUALIFICATION dummy has no effect on the crime equation once we include it in the selection process. Actually, none of the variables used as "instruments" have a significant effect on domestic victimization once they are included in the selection equation.

coefficient is still negative and very significant.[59] Most importantly, there is no support of sample selection bias, as suggested by the estimated correlation coefficients between the errors of the two equations, which are not statistically significant different from zero. Moreover, notice that most of the estimated coefficients of the sample selection models are very similar to the ones from the simple Probit model of specification 2. This suggest that the results of specification 2 can be used.[60]

From the results in specification 2 we can see that the coefficient of immigration status is slightly smaller than in face-to-face interviews. Using the representative individual used in the previous section for domestic violence we find that, the probability of an immigrant to suffer a domestic crime is 2.36%, while the same probability for their native counterparts is 3.92%. Thus, the estimated difference is 1.55 percentage points or a decrease of around 66%, which is lower than the relative effect in face-to-face interviews. However, the difference in the effect of immigration status on victimization between face-to-face and self-completed interviews is not large enough to be interpreted as more under-reporting by immigrants. We might observe this difference just because of the different nature of the self-completion questions, or because of sampling error, or because of differences in the sample size.[61]

6.2.2 The Presence of Others during the Face-to-Face Interview

Presence of other family members during the (mainly face-to-face) interview process may affect the reporting behavior of the respondents (see, for example, Conti and Pudney, 2011), and it may actually result in under-reporting if the questions refer to very sensitive information (see, Acquilino, 1993). Particularly, we expect that the presence of respondent's partner, mainly if the respondent is a female, would reduce the probability of reporting a domestic crime. This might be because of fear of reprisal if the partner is also the offender, or because the respondent does not want to reveal to partner a crime that suffered by other family member, such as parents. As a first indicator, the data show that the probability to report a domestic crime is only 0.19% when the partner is present but 0.57% when the partner is not present, an increase of 200%.

Using this information we are able to say something about the reporting behavior of immigrants relative to the reporting behavior of natives. Particularly, we examine whether the effect of being an immigrant on the risk of victimization in the cases where the partner is present is

[59]Note that more positives help to obtain more precise estimates, but the smaller sample size reduces precision.

[60]We have also used the IPW estimator, constructing the weights as the inverse of the predicted probability to participate in self-completions. These probabilities were estimated by running a Probit model of s on $\mathbf{x}$ and other regressors that are intentionally excluded from $\mathbf{x}$, such as education dummies, language difficulties, etc. The results of the IPW are very similar to the simple Probit ones, which suggest that the simple Probit estimator is also consistent.

[61]Note that a Probit model for face-to-face interviews holding only the sample from specification 2, the immigration coefficient becomes insignificant. However, it is highly likely that this is because of the very small number of positives of the domestic crime variable combined with the relatively smaller sample. Finally, regarding the effects of the other variables on the crime equations from models 2, 3 and 4 we have: risk of victimization decreases with age, London being the least risky place. Concerning the selection equation in specifications 3 and 4 we observe that the probability of selection decreases as age increases. Full results are available from the author on request.

different from this effect in the cases where the partner is not present. According to this, if immigrants under-report by more than natives, the estimated gap will be larger in the cases where partner is present (more negative). This could be formulated using the Probit model below,

$$E(y_i|\mathbf{x}_i) = \Phi(\beta_0 + \beta_1\textsc{Immigrant} + \beta_2\textsc{Par.Present} + \beta_3\textsc{Immigrant} \times \textsc{Par.Present}), \quad (1)$$

where y_i, is the binary variable that takes the value one if a person is victimized by a family member and zero otherwise. The coefficient of the interaction term β_3 is of main interest. Holding everything else constant, if immigrants under-report by more than natives, we expect this coefficient to be negative. Of course, here we also assume that immigrants' reporting behavior does not differ from natives' one under no presence of the partner. Most importantly, this strategy requires that PARTNER'S PRESENCE is assigned randomly, so that people whose partner was present are not different from people whose partner was not present. However, this is not the case. Probit results show that people whose partner was present are relatively more males, less educated, more married, less employed, and stay relatively more in more deprived and urban areas. Also, age has an inverse U-shaped effect on partner's probability to be present. Therefore, a more appropriate strategy would be to examine the differences in reporting behavior between immigrants and natives once we control for all these characteristics.

The main results are presented in Table 14 in three specifications (without controls, after controlling for age, gender, and area dummies, and after in addition controlling for marital, education, and employment status). The results are very interesting. We can see that in specification 1, $\hat{\beta_3}$ (the estimate of β_3) is actually positive and statistically significant, which indicates that if one group under-reports it is the one of natives. Although this estimated coefficient becomes insignificant once we use the previously discussed regressors, it is still positive and preserves some of its magnitude. Particularly, this difference exists due to the following: a) immigrants whose partners are present actually report (insignificantly) more that immigrants whose partners are not present, but b) natives whose partners are present report (insignificantly) less than natives whose partners are not present (but significantly less only in specifications 1 and 2).[62] Thus, this might indicate that immigrants' reporting behavior does not alter in the presence of their partners but natives' one does.

For the above result to take a meaningful interpretation in terms of more under-reporting by natives we conduct two further exercises. Firstly, we examine the reporting behavior in computer-based self-completions, where the presence of the partner should have a much smaller effect both because respondents under-report relatively less in self-completions, and because there were clear instructions by the interviewer that the partner was not allowed to disrupt the interviewee by any means (for instance, not allowed to look at the computer's screen). The results, which are again shown in Table 14, are quite interesting. From specifications 1 and 2 we can see that both immigrants and natives report significantly less crime when partner is present

[62]Differently, immigrants without the presence of their partner report significantly less than natives without the presence of their partner, but immigrants with the presence of their partner report insignificantly more than natives with the presence of their partner.

than in the cases when partner is not present. However, this is due to differences in observed characteristics between people whose partner is present and people whose partner is not present, since in specification 3 it is clear that the reporting behavior of both immigrants and natives does not change under the presence of their partner. Secondly, although the presence of the partner may affect the reporting behavior for domestic crime, it should not affect the reporting behavior for crimes suffered by acquaintances (or, it should affect it by much less). Indeed, the results in the lower parts of Tables 14 are very similar to the results of self-completions. Specification 3 shows that the impact of being an immigrant on the probability to suffer a crime of acquaintances is significantly negative, and not statistically different between the cases where the partner was present and where the partner was not present.

Overall, from both strategies used we can conclude that, there is no evidence that immigrants under-report domestic violence by more than natives and perhaps, immigrants report more accurately than natives. Thus, there is also no reason to believe that immigrants would under-report by more than natives crime suffered by acquaintances either. If we were able to observe the true victimization incidents, the immigration-victimization link could be even larger. Thus, there should be other reasons to explain the above pattern. This is examined in the following two subsections.

6.3 Network Effects and Assimilation Patterns for Violent Crime

As discussed in subsection 5.2.2 a reason why immigrants are less likely than natives to suffer a domestic crime or a crime by acquaintances could just be because immigrants, particularly the most recent ones, have on average fewer family members or a smaller network of acquaintances.

First of all, this cannot be the case for domestic violence, as the distribution of the variable Number of Household Members reveals that (even the most recent) immigrants have actually more members in their families than natives, even if we control for differences in age distribution.[63] Unfortunately, the BCS does not provide any information on the number of respondents' acquaintances. Nevertheless, in this subsection, I examine a **network effect** hypothesis, by assuming that immigrants expand their network of acquaintances as time spent in the country increases. Therefore, the variable Number of Years of an Immigrant in the Country is used, once I control for immigration status and basic demographic characteristics. If a network effect exists, we expect this variable to have a positive significant effect. Note, however, that this variable will also capture other assimilation patterns not related to expansion

[63]Actually, a Poisson regression of the number of household members on immigration dummy and a linear trend for the number of years in the country, and controlling for differences in immigrant-native age distribution (including a cubic on age), shows that being a very recent immigrant (who just entered the country) increases the mean number of family members from 2.28 to 2.39, a difference that is statistically significant at 1%. Moreover, being an extra year in the country adds 0.002 members in a family, which is also significant, but only at 5%. As expected, if we do not control for age differences, being an immigrant increases the family size by almost one person, which is party because immigrants are younger but younger respondents tend to have more household members, but being an extra year in the country decreases the family size by 0.028 members. Both differences are very statistically significant.

of networks and it is very difficult to disentangle these two.

At a first glance, the results presented in Table 15 provide some support on the above hypothesis. We can see from specification 1 that the linear trend has a positive and significant at 10% effect. Thus, more recent immigrants are much less victimized than natives, but immigrants' victimization probability converges to natives' one as time spent in the host country increases, partially due network effects. However, we can also see that this assimilation is very slow as it takes around 70 years for immigrants to reach natives' probability of victimization. Moreover, from specification 2, where a quadratic term is also included, it is clear that starting with a very large difference, the estimated immigrant-victimization gap closes as time spent in the country increases but in a decreasing rate. Actually, a closer investigation shows that the immigration-victimization link never becomes positive, but it goes very close to zero at around "30 years in the country", getting larger and negative once again beyond this point, becoming significant after around "55 years in the country".

Nevertheless, if this assimilation pattern was solely due to network effects, we would expect no assimilation patterns for domestic crime (unless the number of family members increases much more rapidly for immigrants) and crime by strangers. However, from specifications 3 we can see that there is a linear assimilation trend for Domestic Crime as well.[64] Moreover, specification 4 indicates that a weak quadratic assimilation pattern exists for Crime by Strangers too, once we have controlled for racially motivated crime (see next subsection), even though networks should play no role for crime by strangers in this context.

Given the evidence from all four specification, it seems that if network effects exist for Crime by Acquaintances, they are quite weak. According to these results the following story could be more appropriate. More recent immigrants, perhaps because they consider themselves more vulnerable, set strategies associated with lower victimization. As time spent in the country increases, immigrants assimilate in natives' lifestyle, or increase their networks of familiar people, resulting in a smaller victimization difference. However, for earlier immigrants, the picture looks different for Crime by Acquaintances and Crime by Strangers. Earlier immigrants seem to be less likely to suffer violent incidents even though we control for differences in the age distribution.[65] Hence, earlier immigrants, due to some unobserved factors, follow social lifestyles associated with lower victimization than natives with the same basic demographic characteristics.

6.4 Controlling for Racially Motivated Crime

Racially motivated crime (RMC) has been the subject of many monographs, such as Gabbidon and Greene (2009), Spalek (2008) and Kalunta-Crumpton (2010) to mention only a few. Tra-

[64]The quadratic trend does not fit well in domestic crime.

[65]It is essential to control for Age because earlier immigrants are older and therefore, they have a lower victimization probability. Moreover, it is important to stress that controlling for Age by including a quadratic or a cubic term instead of dummies makes no difference in the assimilation patterns found for domestic and crime by acquaintances (but slightly weakens the assimilation relationship for crime by strangers). Therefore, we can argue that it is not the case that we observe the negative relationship for earlier immigrants because we were not able to capture the effect of Age properly.

ditionally associated with ethnic minorities, RMC refers to "hate crime" against individuals of different ethnic group. As opposed to violent crime in general, RMC does not require interactions or interrelations between the potential victims and potential offenders. Offenders, most probably extremists of one race, violently abuse people of a different race, color, or religion, without any pre-existent argument, and in most cases without any provoking action by the victim. Since 43% of immigrant population in the BCS data consists of nonwhite individuals as opposed to only 2.5% for natives, immigrants are more likely to suffer RMC.

Of course, a main difficulty in empirical studies of RMC is to find appropriate data. Moreover, as RMC is traditionally associated with ethnic minorities, occasionally, researchers ignore that white people can also be victims of RMC.

In the present study I deal with RMC as follows. For each victimization incident a question is asked about whether the victims think that the incident was of racial motive. As the question is asked to every victim, we control for RMC against white people as well. However, the problem is that I only observe *perceived* RMC rather than *actual* RMC. Therefore, we need to take into account that, as RMC is traditionally associated with minority groups, ethnic minorities could be more likely to consider a violent crime as being of race motive compared to white individuals, even if the crime is of the same nature. Nevertheless, in this study I assume that victims' *perceived* RMC coincides with *actual* RMC. In the data only 37 victims of violent crime out of 1,190 victims perceived an incident suffered as RMC (3.11%). From these 37 victims 17 were immigrants (around 17% of immigrant victims) and 20 where natives (which is only the 1.83% of native victims).

Thus, I can identify all cases of racially motivated incidents and control for them by replacing their values with zeroes.[66] First of all, from Table 16, we can see that apart from one case, RMCs were committed by strangers, which is consistent with the argument that pre-existed history and interrelations are not needed for this crime to take place. In Table 17, as a first indicator, a simple mean comparison is presented before and after controlling for RMC. We can see that the immigrant-native difference in the probability to suffer a CRIME BY A STRANGER alters sign. However, it is still far from the differences in DOMESTIC CRIME and CRIME BY ACQUAINTANCES.

In the next stage we look at the relationship between IMMIGRANT and CRIME BY STRANGERS once we control for RMC and for basic demographic characteristics. This is examined in Table 18. The second specification presents the estimates for CRIME BY STRANGERS once we replace the cases of RMC with zeroes. Specifications 1, 3 and 4 present the estimates for CRIME BY STRANGERS without controlling for RMC, and the estimates for CRIME BY ACQUAINTANCES and DOMESTIC CRIME for the sake of comparisons. From specification 2 we can see that if RMC did not exist and if immigrants faced the same area, gender and age distribution, they would be less likely to be victimized by strangers as well, a difference that is statistically significant at 10%. The marginal effects, using the representative individual of subsection 5.2.2, show that after controlling for RMC, being an immigrant does reduce the probability of CRIME BY STRANGERS

[66] I have also tried dropping all RMCs from the sample. The results are almost identical.

by 0.49 percentage points (which is significant at 5%), a relative decrease of around 37%, whereas there is no change in the estimated probability of crime if we do not control for RMC.

Thus, we can see that controlling for these few cases of RMC is enough for changing the picture for crime suffered by strangers. However, comparisons with specifications 3 and 4 show that controlling for RMC is not enough to explain the estimated differences between the three crime types. As we can see from specification 2 and 3 the relative effects are much larger for Crime by Acquaintances and Domestic Crime. Thus, although RMC is able to explain some of the unexpected difference in the estimated immigrant-native victimization differentials by relationship status, some unobserved reasons remain.

Summing up, immigrants face a lower probability of violent victimization and we have argued that this might be because immigrants follow social lifestyles associated with lower victimization. However, further analysis showed that this difference exists only for crime by familiar people, as immigrants face the same victimization probability for crime by strangers (if we do not control for racially motivated crime). This can be considered as unexpected if immigrants follow crime-avoiding lifestyles, given that violent crime depends a lot on interactions between potential victims and potential offenders. However, we provided evidence that this difference is not because of more under-reporting by immigrants. Moreover, we showed that there is only weak evidence of network effects for crime by acquaintances. In addition, some of this difference can be explained by the fact that immigrants are more likely to suffer RMC.

But why is the immigration-violent victimization link weaker for crime by strangers than for crime by acquaintances and domestic crime? If we accept that immigrants are more "crime-avoiders" than natives, this can be explained by the fact that it is highly likely that a large proportion of immigrants' "familiar people" consists of immigrants, so that there is an **enhancement effect**. On the contrary, the proportion of immigrants in the pool of strangers that an immigrant meets should be more similar to the proportion of immigrants in the pool of strangers met by a native. Therefore, the enhancement effect does not apply in this case.[67]

Another interesting question emerging from our analysis is the following: if immigrants set the aforementioned lifestyle strategies, why do we not observe a negative association for personal thefts as well? First of all, as has been stressed throughout this paper, personal behavior is a highly more important determinant for violent crime than for personal thefts. Therefore, the aforementioned lifestyles of immigrants would have a much stronger effect on violent crime than on personal thefts, which has as a result to overbalance the positive victimization-immigration association because of higher *proximity* for violent crime, but not for personal thefts. In a cost-

[67]Following the simple calculation in subsection 2.1, assume again that the probability of committing a crime is 6% for an immigrant and 10% for a native. However, now assume that there is 10% probability for a native to interact with a stranger immigrant (which is about the proportion of immigrants in the UK) but there is 25% probability for an immigrant to interact with a stranger immigrant (since it is still more likely that an immigrant will interact with strangers from the same ethnic background due to concentration of immigrants in specific areas.) Thus, according to this simple example, holding everything else constant, the probability for an immigrant to suffer a crime by stranger is $6\% \times 0.25 + 10\% \times 0.75 = 9\%$, but $6\% \times 0.10 + 10\% \times 0.90 = 9.6\%$ for immigrants, a difference of 0.6 percentage points. However, for crime by acquaintances this difference was 2.2 percentage points.

benefit setting, the above can be explained by the fact that it is much more costly (in the sense that it needs much higher effort) to reduce the uncertainty of suffering a personal theft than to reduce the uncertainty of suffering a violent crime.

Thus, all evidence of this empirical study indicates that, indeed, immigrants set strategies that correspond to unobserved differences in routine activities or lifestyle-exposure associated with a lower criminal activity compared to natives.

7 Further Topics

7.1 Interaction Term between of Immigration Status and Ethnic Status or Regional Dummies

So far, we have ignored the fact that there is a great deal of ethnic heterogeneity in immigrant population. It might be that, due to cultural differences, immigrants of different ethnic background may follow different social lifestyles associated with different risks of violent victimization. Moreover, location of immigrants is not randomly assigned. Different locations may attract different types of immigrants, or immigrants located in different places may face different conditions, which in turn may affect the strategies they set with regard to their *social lifestyle-exposure* and *routine activities*.

We first look at the former by including interaction terms between immigration status and ethnic background. The results are presented in Table 19 for all violent crime categories. Note that, although only the coefficients of interest are presented, we use the specification where we control for gender, age dummies, and location characteristics (as in the third specification of Tables 7 and 8). Regarding Total Violence, we find that the results shown in Table 8 (where immigration status has a negative significant at 5% effect on violent victimization) are driven by the differences in victimization experiences between White immigrant-native counterparts, and Chinese & Other immigrant-native counterparts, as there are no differences between the other three ethnic group of immigrant-native counterparts.[68] Finally, comparisons among the different ethnic groups of natives show that only Asian natives suffer lower violence than White natives.

The picture is different if at the different violent crime types by relationship status. It is important to stress that for Domestic Crime and Crime by Strangers we only use interactions between White and Immigrant because, otherwise, there was not enough variation to estimate all coefficients of interest. As far as Crime by Acquaintances is concerned, it is clear that White immigrants still face a lower probability of victimization than White natives (and significant at 5%), but this gap closes for Asians and Chinese & Other ethnic groups.

[68] Comparing a White immigrant with a White native, who are both males, between 26 and 35 years of age and live in average deprived urban areas of East England, we find that being an immigrant decreases the probability of violent victimization by 1.09 percentage points (from 0.0534 to 0.0425, a significant difference at 10% significance level). Moreover, regarding the Chinese & Other ethnic group, we find the for the same representative individual, Chinese & Other immigrants' victimization probability is -0.067 percentage points lower than Chinese & Other natives (from 0.0238 to 0.0908, a significant difference at 5%).

Conversely, the difference even increases in magnitude for BLACK individuals.[69] Note that, although negative, the difference in the probability of victimization by acquaintances between NON-WHITE immigrants and NON-WHITE natives is statistically insignificant.[70] However, the picture is quite different for DOMESTIC CRIME. Contrary to CRIME BY ACQUAINTANCES, the main difference here is observed to be between NON-WHITE natives and NON-WHITE immigrants. NON-WHITE immigrants suffer much less domestic crime than NON-WHITE natives but this gap closes for WHITE people.[71] Finally, note that, there are no statistically significant immigrant-native differences across ethnic groups for CRIME BY STRANGERS. In this study I do not go into further investigation on the rationale of the aforementioned observed relationships but I keep the analysis merely descriptive. Thorough investigation would require a much larger data set as there is a very small proportion of positives for each crime category by relationship status. This analysis is left for future research.

Next, I consider interaction terms between immigration status and regional dummies. First of all, in order to be able to identify all coefficients of interest I group regions in four categories, keeping London as the baseline area.[72] Again, I present only the coefficients of main importance but I also control for gender, age, and other location characteristics. The results are presented in Table 20. Concerning TOTAL VIOLENCE, the results suggest that there are not many differences across regions. Both immigrants in London and immigrants not in London are less likely to be victimized than their native counterparts,[73] but this difference is higher for immigrants of London. However, if we consider the four regional groups separately we find that although the sign on the immigration-victimization relationship is still negative it turns insignificant. The only area for which the difference is still significant is MIDLANDS.[74] Almost the same relationships hold for CRIME BY ACQUAINTANCES, with the only difference that now, the difference between immigrants and natives of London is higher in magnitude and that the difference is also significant for the regional group SOUTH. For DOMESTIC CRIME, the results are very insignificant probably because of the very low number of positives. However, we still find that immigrants not in London are less likely to be victims of domestic crime than natives not in London.[75] Moreover, it is found that immigrants of MIDLANDS suffer less domestic crime than natives of MIDLANDS.[76] Finally, no statistically significant relationships are found for CRIME BY STRANGERS, even after

[69]A Wald test that compares BLACK immigrants to BLACK natives shows that this difference is significant at 5% (p-value of 0.0362).

[70]The coefficient of the difference is -0.194 with a standard error of 0.163.

[71]However, there was not enough variation to further examine this relationship.

[72]These groups are, NORTH (North East, North West and Yorkshire & Humberside, 12,863), MIDLANDS (East Midlands, West Midlands and East of England, 15,973 obs), SOUTH (South East and South West, 10,142 obs) and WALES (4,243 obs).

[73]Running a regression of TOTAL ASSAULT on the dummy LONDON, its interaction with immigration status, and the rest of the characteristics, shows that immigrant not from London also face lower crime than immigrants not from London.

[74]This is the case perhaps because it is the region with the highest number of observations. A Wald test of the difference gives a p-value of 0.0348. For South, the Wald test gives a p-value of 0.115.

[75]A regression of DOMESTIC CRIME on the dummy LONDON, its interaction with immigration status, and the rest of the characteristics, shows that immigrant not from London face lower crime than natives not from London with a coefficient of -0.240 which is statistically significant at 5%.

[76]The Wald statistic has a p-value of 0.03.

controlling for racially motivated crime.[77]

7.2 Seriousness of Crime

So far, we have found that immigrants face a lower probability of violent victimization but a similar probability of property victimization compared to natives who exhibit the same basic demographic characteristics. However, we have not referred at all to the seriousness of crimes they have suffered. In this subsection I exploit information from the Victim Forms, where all victims were able to rank the "seriousness" of the crimes they suffered in a scale from 1 (not serious) to 20 (very serious). Since each victim could take up to six victim forms, I averaged the "seriousness" score for each victim and then I created an ordinal variable that takes value '1' if victims believed that the "seriousness" of the crimes experienced is between 1 and 5 (*Not Serious*), '2' if it is between 6 and 10 (*Relatively Serious*), '3' if it is between 11 and 15 (*Serious*) and '4' if is between 16 and 20 (*Very Serious*).[78]

It is clear that since this is a measure of perceived "seriousness", the coefficient of immigration status would be upward biased if for some reasons immigrants tend to perceive incidents of the same actual seriousness as being more serious (and *vice versa*). The results for total crime are presented on Table 21. Specification 1 uses no controls, while in specification 2, in line with previous regressions, we control for basic demographics and in specification 3 we include further controls that might be associated with perceived seriousness. What we find is that, regardless the controls we use (look at specifications 1 to 3), immigrants strongly believe that crimes they suffer are much more serious than what natives believe. As we mentioned before, this might not indicate that immigrants are recipients of more serious crimes if they, for some unobserved reasons, tend to overvalue seriousness relative to natives. Using the cut point estimates and the estimated coefficients from specification 2, we predict that being an immigrant victim:[79] 1) decreases the probability for an experienced crime to be considered as *Not Serious* (1-5) by 8.3 percentage points, a relative effect of 11%, but 2) increases the probability for a crime to be considered as *Relatively Serious* (6-10) by 5.3 percentage points, a relative increase of 26%, 3) increases the probability for a crime to be considered as *Serious* (11-15) by 2.3 percentage

[77]Finally, note that further interaction terms exercises show that the highest differences between immigrants and natives exist for, residents of the most deprived areas (although the effect of the interaction term is statistically insignificant), people who rent, people who are less educated, and single individuals (but only for domestic crime). Moreover, there are no interaction effects between gender and immigration status, apart from crime by strangers (once we control for racially motivated crime) for which we find that immigrant males suffer less crime than native male (significant at 5%), but this difference disapears for females. In general, it seems that the highest differences exist for the most vulnerable groups of immigrants. Perhaps immigrants who believe that they are in weak positions (lower *guardianship* or higher *proximity*) are more in fear of a potential crime against them and therefore, decide to balance their position by exhibiting lower *exposure*. As the findings indicate, the result is to suffer lower crime than their native counterparts, perhaps because the effect of lower *exposure* overbalances the effect of higher *proximity* and lower *guardianship*. This subject is left for future research.

[78]Note that from the 11,208 victims, 66% believed that the victimization incidents they experienced are of seriousness from 1 to 5, 25% from 6 to 10, 7% from 11 to 15 and only 2% from 16 to 20. Moreover, note that creating an ordinal variable with 8 categories gives very similar results.

[79]For these predictions we use the representative individual who is a male between 25 and 35 years old, and live in an average deprived urban area in the East of England.

points, but with a relative effect of 56% and finally, 4) increases the probability for a crime to be considered as *Very Serious* (16-20) by 0.7 percentage points, which account for an even higher relative effect of 89%. Note that all these differences are statistically significant at 1%.

Moreover, specification 4 shows that the immigration status dummy has no effect if we control for ethnic background (but it is still significant at 5% if we include the ethnicity dummies on specification 3). However, specification 5, where we interact immigration status with ethnicity dummies, provides further interesting insights. Although White immigrants do not perceive crimes they suffer as more serious than their native counterparts, non-White immigrants do. Actually, a regression where we only interact immigration status with ethnic group White shows that being a non-White immigrant significantly increases the probability to perceive a crime as more serious relative to a non-White native.[80] In more detail, specification 5 shows that apart from White and Mixed ethnic groups, Black, Asian and Chinese & Other immigrants value their crime experiences as more serious than their native counterparts, although the estimated difference is statistically significant (at 5% significance level) only for the group of Asian people. Finally, further analysis where we look at household crime, personal theft and violence separately shows that the above negative relationship holds for each crime category, but it is slighly less significant for personal crime.[81]

8 Count Data Models

All previous analysis concerned the conditional probability of victimization and provided robust results regarding the difference in the probability of victimization between immigrants and natives across the different crime types. However, the count nature of the victimization variable was totally neglected. Considering the count form of the crime variables and utilizing several count data models could provide some further insights on the determinants of victimization in general, and particularly, on the immigration-victimization relationship. For instance, even though immigrants face a lower probability of violent victimization, the implications of our analysis would be very different if, as will be discussed further below, immigrants experience a higher number of crime incidents than natives.

Count data are directly related to the problem of **repeat victimization**, as someone is said to be a **repeatedly** victimized if he/she has suffered more than one incident of the same crime type within the reference period.[82] Together with the causes of a single crime incident, the understanding of the channels through which repeat victimization occurs has also received a lot of

[80]The coefficient is 0.292 with a standard error of 0.082.

[81]It is actually significant at 10% for violence, but insignificant for personal theft. However, notice that we only have 1,186 cases of violence and 745 cases of personal theft. Full results are available from the author on request.

[82]In general criminologists distinguish between the term "repeated victim" and the term "multiple victim" (see, for example, Hope at al, 2001). A person is a multiple victim if he/she suffered more than one type of crimes in the reference period, regardless of the number of crimes of the same crime type. However, other studies do not distinguish between these two terms (see, Farrell, 1992). In the present study, the victim is said to be "repeated" if he/she suffered more than one crime, of a particular crime-type, within the reference period of 12 months prior to the interview.

attention by criminologists, in an attempt to find alternative effective policies for crime reduction which would in turn allow policy makers to efficiently allocate scarce resources in the areas where people or households face the greatest risks (see, Farrell, 1992, Farrell, Phillips and Pease, 1995, and Osborn at al, 1996). This is important, as crime is found to be concentrated among a small group of people and areas (see, Spelman, 1995, Ellingworth, Farrell and Pease, 1995) and because prior victimization is found to be a very strong predictor of future victimization (Hindelang, Gottfredson and Garofalo, 1978, Ellingworth at al, 1997, and Wittebrood and Nieuwbeerta, 2000). Several researchers have attempted to understand the process of repeat victimization by using count data models (see, for example, Nelson, 1980, Tseloni and Pease, 2003, 2004).

There are reasons to believe that the data generating process of zeroes and ones is to some extent different from the process of generating the number of victimization incidents. If this is true, the effects of the characteristics associated with victimization, are allowed to be different between the zero-one process and the number of incidents given at least one victimization incident. First of all, criminologists have made some effort towards understanding whether there is some kind of **state-dependence** among crimes suffered by the same individuals, or it is just that the characteristics associated with higher risks responsible for a first crime incident persist over time resulting in further actions against them. State-dependence among sequences of crimes against the same individuals could be possible if a first crime initiates a positive or a negative "contagious" process. For example, a positive "contagion" (mostly for household crimes) could be the consequence of some kind of transmission of information amongst offenders concerning the vulnerability or *attractiveness* of some targets. Differently, a positive "contagion" for violent victimization could exist if following the victimization incident, victims choose to revenge or retaliate, which in turn would expose the victim to further violence. On the other hand, negative "contagion" would be the result of reevaluation of strategies following an incident, which would make victims, for instance, to increase their *guardianship* or reduce their *exposure*. However, we need to stress that these dynamics cannot be identified in the absence of panel data, which is the case in the present study. This is because, firstly, we are not able to observe when the first action occurred, and secondly because the cross-section models used in this study assume independency of the incidents.[83]

In addition, the effect of the regressors could be different between the two processes (victimization or not, and the number of incidents conditional on victimization), if there is unobserved heterogeneity within the same variables which is associated with differential victimization across the two processes. As an example, consider the relationship between gender and violence. As males exhibit a much higher *exposure*, the victimization probability is much higher for them. However, the picture is different if we consider only victims, as females may be victimized more frequently, perhaps because of domestic violence. A similar story can be considered for im-

[83]We need to note that although the Negative Binomial distribution is consistent with a count generation process with positive "contagion", in the absence of panel data we cannot distinguish between "contagion" and "heterogeneity" among different groups. For interesting details on the genesis and other aspects of the Negative Binomial distribution the reader may refer to Johnson, Kemp and Kotz (2005), Cameron and Trivedi (1998), and Winkelmann (2008).

migrants. Given that immigrants are generally more vulnerable (lower risk for offenders) they decide to set strategies associated with lower *exposure* to crime. As a result, according to the findings of the previous sections, immigrants are on average less likely to be victims of violence. However, if we consider the population of the victims only, it might be the case that here we have either the immigrants that failed to successfully set the low *exposure* strategies, or groups of immigrants whose cultural characteristics are associated with higher *exposure* relative to the groups of less victimized immigrants. According to this, immigrant victims could be equally or even more victimized than native victims. Therefore, if the above were true, we would expect that the effect of immigration status on victimization would be smaller in absolute value in the count data models than in binary models, as the number of the incidents is also taken into account.

In this study I consider both Poisson and Negative Binomial 2 models, as the latter also takes into account over-dispersion (by allowing for an unobserved gamma distributed error) which, as described before, is evident in victimization data.[84] However, the nature of victimization data gives rise to two issues that require special attention. Firstly, as crime is a rare event (at least if we want to consider the different crime types separately), the number of zeroes is very large. Moreover, the (very few) positives are quite dispersed for most of the crime categories. This has harmful consequences on the robustness of the count data estimators. Secondly, there are few cases of victims that reported extreme number of crimes.[85] Count data models are very sensitive to these cases, particularly when the positive counts are too few to identify the parameters of the variables assumed to affect the mean.

I deal with the second issue using two different approaches. Firstly, I censor the crime variable at different points of the violent crime distribution and then I use a Poisson model with a modified likelihood function that takes into account the censoring in the dependent variable (for details in censored count data, see, Terza, 1985, and Brannas, 1992). Not only does this strategy tests for the robustness of the estimates of the conventional count data models, but it also adds some robustness to the estimator.[86]

Furthermore, I use the "Quantile Estimator for counts" developed by Machado and Santos Silva (2005) and successfully applied by several researchers, including Winkelmann (2006), and Miranda (2008). It is well known that the usual quantile estimators are developed for continuous data (see, Koenker and Bassett, 1978) and are not available for counts, or other discrete choice variables. However, Machado and Santos Silva (2005) suggest a method that overcome this problem by adding a uniformly distributed noise to the count outcome (a method called "jittering"),

[84] It is well known that the Poisson distribution assumes equi-dispersion meaning that the first two moments are equal to each other. For more details on count data models refer to Winkelmann (2008) and Cameron and Trivedi (1998).

[85] The maximum they could report in each victim form was 97 crimes. Therefore, in the extreme, someone could report 582 crimes.

[86] On the other hand, the results of censored count count data models are at same time less "robust", because the results of the Poisson Pseudo Maximum Likelihood (see, Gourieroux, Monfort, and Trognon, 1984) do not hold. Thus, the censored model would be misspesified, if the remaining counts above the considered cut-off point do not follow the poisson distribution. In the contrary, the Poisson distribution only requires correct specification of the conditional mean, with valid inference given by the Quasi Maximum Likelihood standard errors.

which artificially generates the required smoothness of a continuous variable. Then, quantile estimation proceeds by using standard quantile techniques. Moreover, they propose averaging out the uniformly distributed noise by considering m jittering samples, a method that which increases efficiency of the estimator.[87] Utilization of this estimator serves two purposes. Firstly, the quantiles are insensitive to the extreme cases, and secondly, we can estimate the effect of the regressors on different parts of the distribution (which might be different according to the repeat victimization theories). However, as the number of zeroes in my dependent variables is very high, it is more sensible to look at the effects of the variables on very high quantiles.

Alternatively, a model that would be also consistent with the story of differential repeat victimization could be a hurdle (two-part) model, where the "hurdle" is set at no crimes. According to this model (see, Mullahy, 1986), zeroes or positives (without distinction on the number of incidents) are generated by a mechanism appropriate for binary choice models. If the realization is positive, the hurdle is crossed, and positives are generated by a truncated at zero (see, Grogger and Carson, 1991, and Gurmu, 1991) distribution for counts, such as the truncated at zero Poisson or the truncated at zero Negative Binomial distribution.[88] Hence, this model explicitly allows to separately model the binary outcome (victimized or not) from the positives (number of incidents given victimization, or differently, repeat victimization).[89] Therefore, we can directly observe whether the independent variables have different effects below and above the hurdle (thus, at different parts of the distribution). Again, the very low number of positives and the extreme reports by some individuals will constraint my analysis. Nevertheless, as an alternative and in line with the Censored-Poisson model, I develop a two-part model for censored counts. Details on the probability and likelihood functions of this modified Hurdle-Censored Poisson model are presented in the Appendix.

In this study I only present results on violent victimization, which was the centre of attention in the main analysis. Furthermore, because of the aforementioned large number of zeroes, the estimation analysis is not so reliable if we further decompose violent crime by relationship type, particularly when I use the Hurdle-Censored Poisson estimator. Thus, I mainly present results on total violent victimization and I refer to results of the separate groups when necessary.[90] Before presenting the count data results, the complete distribution together with the three unconditional moments of the violent crime variables are presented in Table 22. It is clear that TOTAL VIOLENCE is a rare event as only 3.54% of respondents reported at least one violent incidence. It is also clear that incidents of violence are highly dispersed and skewed to the right, a feature driven by DOMESTIC CRIME and CRIME BY ACQUAINTANCES, as CRIME BY STRANGERS is generally concentrated on the first 10 counts.

The results of the count models are presented in Tables 23, 24 and 25. The second and

[87]For details on this estimator refer to Machado and Santos Silva (2005).

[88]Count hurdle models (together with some modified count hurdle models) are very successfully used in health economic literature (see, for example Pohlmeier, Ulrish, 1995, and Gurmu, 1997), or other contexts (see, for example, Gurmu and Trivedi, 1996, Santos Silva and Covas 2000, and Helstrom, 2006).

[89]By taking into account the different data generating process below the hurdle and above the hurdle, we explicitly take into account the exceptional nature of zeroes.

[90]All results that are not present here are available from the author upon request.

the third specifications of Table 23 depict the results of the conventional Poisson and Negative Binomial 2 (NB2) regression models, whereas specification 1 gives simple Logit results for the sake of comparisons. The rest of the specifications present the Censored-Poisson model, where we censor the dependent variable at 5, 10, 15, 20 and 25 crimes. Table 24 displays the estimates resulting from the Quantile estimator for counts, where we look at the effect of the variables at the 25th, 50th, 70th, 80th, 90th, 95th, 99th, 99.9th, and 99.99th percentiles of the distribution. We explore unusually high quantiles because, as I mentioned earlier, it is important to explore the impact of the regressors on the very right end of the distribution. It is also important to note that my results are obtained using 100 jittered samples. Finally, Table 25 shows the results from a simple Hurdle-Poisson model and the results from the modified Hurdle-Censored Poisson model, where I censor at 5 and 10 crimes.[91] The first column gives the probability of crossing the hurdle for which the Poisson distribution is also used,[92] the second specification shows the Zero-Truncated results without censoring and the rest of the specifications provide the findings of the Zero-Truncated Censored models.

To begin with, apart from the coefficient on URBAN and the fact that regional dummies have a smaller effect (relative to London) in NB2, the results of Poisson and NB2 are fairly similar. According to the NB2 model, there is quite strong evidence in favor of over-dispersion.[93] However, it is important to stress that although the Poisson regression model assumes equi-dispersion (conditional mean equal to conditional variance), which implies that the variance-covariance matrix is misspecified under the presence of over-dispersion, it is absolutely valid even in the cases of very over-dispersed data. This is because, as the results of the Pseudo-Maximum Likelihood show (see, Gourieroux, Monfort and Trognon, 1984), the Poisson Maximum Likelihood Estimator consistently estimates the conditional mean and valid inference for the variance-covariance matrix of the estimator is obtained by using Quasi Maximum Likelihood standard errors.[94]

Comparing the binary information with the conventional count data models, several interesting points emerge that need some discussion. Firstly, we can see that although the IMMIGRANT coefficient is still negative, it is now insignificant. However, this should not be interpreted as higher repeat victimization of immigrants without further investigation. Indeed, the Censored-Poisson models, regardless of the cut-off point of censoring, show that the effect of immigration is still very significant and not much different in magnitude than when we use the binary information only. This suggests that the long right tail of the observed distribution affects the precision of the effect of IMMIGRANT on TOTAL VIOLENCE. The results of the Quantile es-

[91] The estimation procedure was numerically unstable when censoring at higher than 10 crimes was considered. This is probably because of the small number of observations above the hurdle (1,190 observations). Therefore, the results of censoring the variable at a higher value are not presented here.

[92] These estimates are closely comparable to Logit ones. Actually, the Logit probability function can be also obtained from considering only the zero probability from the Geometric version of the Negative Binomial distribution for count data (see, Mullahy, 1986).

[93] In this table, $\hat{\alpha}$ is the estimated variance of the gamma distributed unobserved effect. According to the NB2 model the conditional variance of the dependent variable is given by $\omega = \lambda + \alpha\lambda^2$. As the estimated $\hat{\alpha}$ is around 40 and statistically significant, the variance is much higher than the mean.

[94] However, note that if the true data are truly generated by a Negative Binomial distribution, Poisson Pseudo Maximum Likelihood is less efficient than a Negative Binomial model.

timator and the Hurdle-Censored Poisson model are relatively in line with the aforementioned analysis. Regarding the Quantile Estimator results, although the immigration dummy has no effect on the first quartile and the median, as expected due to the small number of positives, its effect is negative and significant along the right part of the distribution. It is also clear that the effect starts diminishing when considering very high quantiles. Finally, from Table 25, the zero-truncated but uncensored Poisson assigns a positive but very imprecisely estimated coefficient to the immigration dummy which, however, turns negative in Zero-Truncated Censored models if we censor at 10 crimes. Overall, these results indicate that the very few observations at the end of the observed distribution reduce the influence of the immigration dummy. This suggests that immigrants may be more repeatedly victimized than natives only for individuals who suffer a large number of incidents. Unfortunately, the sample size does not permit further investigation and safer conclusions.

The most striking result is that although being a male increases the probability of suffering a crime, it actually decreases the mean number of crimes. The Hurdle-Poisson models present this picture clearly. Conditional on being victimized, being a male significantly decreases the number of incidents. The effect is smaller for the Zero-Truncated Censored models but still very significant. Further investigation shows that this result is primarily driven by the relationship between gender and DOMESTIC CRIME. Nevertheless, a negative relationship holds for CRIME BY ACQUAINTANCES too, although it is less significant, but not for CRIME BY STRANGERS. Thus, there is some evidence that although males are more exposed on the incident of violence, females are more repeatedly victimized. This is probably because some women are captured in "unhealthy" relationships that bring them in situations of a constant high risk of victimization.

Finally, it is also interesting that although the effect of URBAN is positive but insignificant in the binary models, it turns negative in the count models. From the Quantile regressions we can observe that the effect of URBAN is the highest between the 90th and the 95th percentiles and then decreases turning negative after the 99th percentile. Similar conclusions are obtained from examining the Hurdle-Censored Poisson model. Further analysis shows that this result is driven by the impact of being in an urban area on CRIME BY STRANGERS.[95] Even though people in urban areas face a significantly higher risk of victimization by strangers, repeat victimization by strangers is higher in rural areas if we only consider victimized individuals. This indicates that in rural areas there is a higher concentration of CRIME BY STRANGERS among the same individuals compared to urban areas. This is an interesting finding, but further research is required to identify the reasons behind this relationship.[96]

[95] Being in urban areas significantly increases the victimization risk, where the estimated coefficient of URBAN in the Logit model is 0.358 with a p-value of 0.005. However, in the Zero-Truncated Poisson this estimate is negative (-0.647) with a p-value of 0.014.

[96] Further investigation of these models with regard to the effects of the other variables can result in many interesting implications. However, as this paper concentrates on the victimization-immigration relationship this analysis is skipped here, but it is subject to future research.

9 Conclusion

This study presented a comprehensive analysis of the relationship between immigration status and victimization in England and Wales using the 2007/08 sweep of the British Crime Survey.

Initially, we presented some evidence on the immigration-victimization relationship for Inside and Outside Burglaries. Immigrants' households are more at risk of Inside Burglaries but this is well explained by the fact that immigrants reside relatively more than natives in urban and more deprived areas where the incident of an Inside Burglary is highly more likely. On the other hand, a negative relationship was found between immigrants' households and the incident of Outside Burglaries. We argued that this is probably because immigrants own a smaller amount of properties that are subject to Outside Burglaries such as, outhouses, garages, etc. This argument was supported with results on assimilation patterns (earlier immigrants are better settled and therefore own more outside properties than more recent immigrants) and zero-inflation count models (which show that a higher proportion of immigrants belong to the zero inflation category, meaning that immigrants are in lower risk just because they own fewer outside properties).

Furthermore, we showed evidence on Personal Thefts, a crime that is of a very different nature since, although instrumental as well, it entails personal contact. The results indicated that immigrants are in higher risk of Personal Thefts, but most of this positive association can be attributed to the fact that they disproportionately reside in the areas of London where the incident of a Personal Theft is much more likely than the other regions in England and Wales.

Next we presented a series of evidence for Violent Crime, in which this work focuses on. Violent Crime, as opposed to the aforementioned categories, is an expressive type of crime where interrelations and interactions between the potential victims and potential offenders are vital. According to this, personal behavior is a much stronger predictor for Violent Crime than for Personal Thefts and Household Crime. Even after controlling for a large set of characteristics associated with violent victimization, the empirical analysis indicated that immigrants still face a significantly lower risk of violence. A possible explanation, which relies on the theoretical views of this paper, is that immigrants set strategies (that determine their *lifestyle-exposure* and *routine activities*) which are associated with a lower risk of violent victimization relative to natives. Nevertheless, a closer examination indicated that the negative association is due to the lower risk of victimization by Acquaintances and lower risk of Domestic Crime, since the regression results showed that there is not any association between being an immigrant and crime suffered by Strangers. This result is, at a first glance, not in line with the hypothesis mentioned above, since if immigrants follow a particular lifestyle associated with lower *exposure* and therefore, lower crime, we expected to observe a negative association for crime by strangers as well. Thus, the next section attempted to shed light on the differences in the estimated immigration-victimization associations across the three (by relationship status) Violent Crime types.

Firstly, we examined the reporting behavior of respondents towards Domestic Crime, as there is evidence that respondents tend to under-report domestic crime in face-to-face interviews. Thus, if immigrants tend to under-report crime suffered by family members by more than natives, the observed immigration-victimization association will be downward biased. However, both strategies that we followed showed no evidence that immigrants under-report Domestic Crime by more than natives, and therefore, there is no reason to believe that they would under-report crime by Acquaintances either. Particularly, in the first strategy we used data on computer-based self-reported crime, as there is evidence that people respond much more truthfully in computer-based than in face-to-face interviews. The results from computer-based interviews are in line with the results from face-to-face interviews, that is, immigrants are significantly less likely to be victims of domestic violence. In the second strategy, we explored the information on whether respondents' partners were present during the face-to-face interviews, as people may under-report domestic crime by more in the presence of their partner. After a thorough analysis, also comparing with results in crime by acquaintances and computer-based self-reports of domestic crime, we concluded that if one group under-reports, this is the group of natives.

In the second step, we tested whether the lower risk of victimization by Acquaintances that immigrants face is just because of the fact that the groups of acquaintances are relatively smaller for more recent immigrants. Therefore, connecting that to assimilation patterns, we showed that more recent immigrants are actually in lower risk of victimization than earlier ones. However, showing some further evidence, we argued that the observed assimilation link was most probably driven by other unobserved assimilation features. If network effects exist, they are relatively weak, and by no means could they explain the observed differences across violence crime types. Last, we investigated what would be the effect of being an immigrant on crime suffered by strangers if RMC did not exist. Interestingly, we showed that if we control for (the only 37 cases - 20 for natives and 17 for immigrants - of perceived) RMC, which is a much more "random" crime highly associated with ethnic minorities, the risk of suffering a Violent Crime by Strangers becomes negative and significant at 10%, but not of the magnitude observed for crime by acquaintances people and family members. The **enhancement** effect is a possible explanation for this. It is highly likely that a large proportion of immigrants' family members and acquaintances are immigrants as well. If immigrants follow a crime-avoiding behavior, then the low victimization of an immigrant by his family and acquaintance group happens both because: 1) they are crime-avoider themselves; 2) the reference group is composed in large part of crime-avoiders. On the contrary, the pool of strangers that an immigrant meets is composed by the same proportion of immigrants as the pool of strangers met by a native. Therefore, the enhancement effect does not apply in this case.

Next, we briefly discussed the seriousness of the crimes that victims face. We actually found that although immigrants are less likely to be victims of violent activity, they consider the crimes they suffer as more serious than the crimes natives suffer. Of course, if for any reasons immigrants tend to perceive crime of the same actual seriousness as more serious, all results of this section are biased upwards. Moreover, a very brief analysis of decomposition of immigrants by ethnic status and location did not reveal any important relationships. However, a much closer examination is required, perhaps using even larger data sets (by pooling several sweeps from the British Crime Survey).

After establishing the above relationship for the probability of victimization, we considered count data models, exploiting the count nature of the VIOLENT CRIME variable. Count data analysis is important because it is directly connected with the concept of repeat victimization. As explained in detail in Section 8, the implications of our analysis would be very different if immigrants were disproportionately victims of repeat crimes. Several models were considered (Poisson, NB2, Censored Poisson, Quantile Estimator for counts, Hurdle-Censored Poisson) to explore the association between the number of violent victimization incidents and immigration. Initially, conventional Poisson and NB2 models showed that once we take the count information into account, immigrant coefficient loses much of its significance and magnitude. However, this should not be interpreted as differential repeat victimization by immigrants, as the Censored(-Hurdle) Poisson, and the "Quantile for counts" estimator showed that this result was driven by the very end of violent crime distribution. This means that if differential repeat victimization between immigrants and natives exists, it does only among highly victimized individuals. Therefore, according to these results, the effect of being an immigrant on violent victimization is relatively similar in both, the probability of suffering a crime and the number of crimes suffered. However, data limitations (very few and dispersed positives) did not allow us to examine the above relationships by the three different violent crime types.

Nevertheless, the use of the count information together with appropriate count data models is very promising and it can provide many interesting insights not only about the relationship between immigration and victimization but also about the determinants of victimization is general. For instance, we showed evidence that the victimization probability is higher for males because of their higher *exposure*, but once we consider the victimized individuals only, females are victimized much more frequently perhaps due to repeat domestic violence. Further analysis is subject to future research, perhaps considering pooling several sweeps from the British Crime Survey in order to increase the sample size, and consequently, the robustness of the estimated relationships.

References

Algan, Y., C. Dustmann, A. Glitz, and A. Manning (2010), "The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom", *The Economic Journal*, 120, F4-F30.

Aquilino, W. S. (1993), "Effects of Spouse Presence During the Interview on Survey Responses Concerning Marriage", *Public Opinion Quarterly*, 57, 358 - 376.

Becker, G. S. (1968), "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, 76, 169-217.

Bell, B., S. Machin, and F. Fasani (2010), "Crime and Immigration: Evidence from Large Immigration Waves," LSE, CEP Discussion Paper No 984.

Bolling, K., C. Grant, and J. Donovan (2008), 2007-08 British Crime Survey (England and Wales), Technical Report I-II, Home Office.

Brannas K. (1992), "Limited Dependent Poisson Regression", *Journal of the Royal Statistical Society, Series D (The Statistician)*, 41, 413-423.

Cappellari L., and S. T. Jenkins (2003), "Multivariate Probit Regression using Simulated Maximum Likelihood", *The Stata Journal*, 3, 278-294.

Cameron, A. C., and P.K. Trivedi (1998), *Regression Analysis of Count Data*, New York, Cambridge University Press.

Cohen, L. E. and M. Felson (1979), "Social Change and Crime Rate Trends: A Routine Activity Approach", *American Sociological Review*, 44, 588-608.

Conti, G., and S. Pudney (2011), "Survey Design and the Analysis of Satisfaction", *Review of Economics and Statistics*, 93,1087-1093.

Curtis, A. L. (1974), "Victim Precipitation and Violent Crime ", *Social Problems*, 21, 594-605.

Deadman, D., and Z. MacDonald (1997), "Offenders as Victims of Crime?: an Investigation into the Relationship between Criminal Behaviour and victimisation", *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 167, 53-67.

Edwards, J. H. and A. W. F. Edwards (1984), "Approximating the Tetrachoric Correlation Coeff-cient", *Biometrics*, 40, 563.

Ehrlich, I. (1973), "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation", *Journal of Political Economy*, 81, 521-565.

Ellingworth, D., G. Farrell, and K. Pease (1995), "A Victim Is a Victim. Is a Victim? Chronic Victimization in Four Sweeps of the British Crime Survey", *British Journal of Criminology*, 35, 360-365.

Ellingworth, D., T. Hope, D. R. Osborn, A. Trickett and K. Pease (1997), "Prior Victimization and Crime Risk", *International Journal of Risk, Security, and Crime Prevention*, 2, 201-214.

Farrell, G. (1992), "Multiple Victimization: Its Extent and Significance", *International Review of Victimology*, 2, 85-102.

Farell, G., C. Phillips, and K. Pease (1995), "Like Taking Candy. Why Does Repeat Victimization Occur?" *British Journal of Criminology*, 35, 384-399.

Felson, B. R., S. F. Messner, A. W. Hosking, and G. Deane (2006), "Reasons for Reporting and not Reporting Domestic Violence to the Police", *Criminology*, 40, 617-648.

Greene, H. W. (2011), *Econometric Analysis*, 7th Ed., Prentice Hall, Pearson Education International.

Grogger, J. T., and R. T. Carson (1991), "Models for Truncated Counts", *Journal of Applied Econometrics*, 6, 225-238.

Gourieroux, C., A. Monfort, and A. Trognon (1984), "Pseudo Maximum Likelihood Methods: Theory", *Econometrica*, 52, 681-700.

Gurmu, S. (1991), "Tests for Detecting Overdispersion in the Positive Poisson Regression Model", *Journal of Business and Economic Statistics*, 9, 215-222.

Gurmu, S. (1997), "Semi-Parametric Estimation of Hurdle Regression Models With an Application to Medicaid Utilization", *Journal of Applied Econometrics*, 12, 225-242.

Gurmu S., and P. K. Trivedi (1996), "Excess Zeros in Count Models for Recreational Trips", *Journal of Business and Economic Statistics*, 14, 469-477.

Hausman, J. A., J. Abrevaya, and F. M. Scott-Morton (1998), "Misclassification of the Dependent Variable in a Discrete-Response Setting", *Journal of Econometrics*, 87, 239-269.

Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47, 153-161.

Hellstrom, J. (2005), "A Bivariate Count Data Model for Household Tourism Demand", *Journal of Applied Econometrics*, 21, 213-226.

Hindelang, M., M. Gottfredson, and J. Garofalo (1978), Victims of Personal Crime: An Empirical Foundation for the Theory of Personal Victimization, Cambridge, MA: Ballinger.

Hope, T., J. Bryan, A. Trickett, and D. R. Osborne (2001), "The Phenomena of Multiple Victimization: The relationship between Personal and Property Crime Risk", *British Journal of Criminology*, 595-617.

Johnson, N.L., A.W. Kemp, and S. Kotz (2005), *Univariate Discrete Distributions*, 3rd ed., New Jersey, John Wiley & Sons.

Kalunta-Crumpton, A. (2010), Race, Crime and Criminal Justice: International Perspectives, edited by Anita Kalunta-Crumpton, Palgrave Macmillan, New York.

Koenker, R., and G. Bassett (1978), "Regression Quantiles", *Econometrica*, 46, 33-50.

Lambert, D. (1992), "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing", *Technometrics*, 34, 1-14.

Lauritsen, J. L., R. J. Sampson, and J. H. Laub (1991), "The Link Between Offending and Victimization Among Adolescents", *Criminology*, 29, 265-292.

Machado, J. A. F., and J. M. C. Santos Silva (2005), "Quantiles for Counts", *Journal of the American Statistical Association*, 100, 1226-1237.

Meier, R. F. and T. D. Miethe (1993), "Understanding Theories of Criminal Victimization", *Crime and Justice*, 17, 459-499.

Miethe, D. T., R. F. Meier (1990), "Opportunity, Choice, and Criminal Victimization: A Test of a Theoretical Model", *Journal of Research in Crime & Delinquency*, 27, 243-266.

Miethe, D. T., M. Stafford, and J. S. Long (1987), "Social Differentiation in Criminal Victimization: A test of Routine Activities/Lifestyle Theories", *American Sociological Review*, 52, 184-194.

Miranda, A. (2008), "Planned Fertility and Family Background: A Quantile Regression for Counts Analysis". *Journal of Population Economics*, 21, 67-81.

Mullahy, J. (1986), "Specification and Testing of Some Modified Count Data Models", *Journal of Econometrics*, 33, 341-365.

Nelson, J. F. (1980), "Multiple Victimization in American Cities: A Statistical Analysis of Rare Events", *American Journal of Sociology*, 85, 870-91.

Noble, M., D. McLennan, K. Wilkinson, A. Whitworth, H. Barnes, and C. Dibben (2008), "The English Indices of Deprivation 2007", Communities and Local Government: London, Crown Copyright.

Osborn, R. D., D. Ellingworth, T. Hope., and A. Trickett (1996), "Are Repeatedly Victimized Households Different?", *Journal of Quantitative Criminology*, 12, 223-245.

Papadopoulos, G. (2013), "Immigration and Criminal Behaviour", University of East Anglia Applied and Financial Economics Working Paper Series, School of Economics, University of East Anglia, Working Paper No. 0 37.

Papadopoulos, G. and J.M.C. Santos Silva (2008), "Identification Issues in Models for Underreported Counts", Department of Economics, Essex University, Discussion Paper Series, No. 657.

Parks, R. (1967), "Efficient Estimation of a System of Regression Equations when Disturbances are both Serially and Contemporaneously Correlated", *The Journal of the American Statistical Association*, 62, 500-509.

Santos Silva, J. M. C., and F. Covas (2000), "A Modified Hurdle Model for Completed Fertility", *Journal of Population Economics*, 13, 173-188.

Schultz, L. (1968), "The Victim-offender Relationship ", *Crime and Delinquency*, 4, 135-141.

Smith, A. D. and G. R. Jarjoura (1989), "Household Characteristics, Neighborhood Composition and Victimization Risk", *Social Forces*, 68, 621-640.

Spalek, B. (2008), *Ethnicity and Crime: A Reader*, Edited by Basia Spalek in Sandra Walklate (ed): Readings in Criminology and Criminal Justice Series, Open University Press, England.

Spelman, W. (1995), "Once Bitten, Then What? Cross-sectional and Time-course Explanations of Repeat Victimization", *British Journal of Criminology*, 35, 366-383.

Terracol, A. (2002), "Triprobit and the GHK Simulator: A Short Note", mimeo.

Terza, V. J (1985), "A Tobit-type Estimator for the Censored Poisson Regression Model", *Economics Letters*, 18, 361-365.

Tseloni, A. (2006), "Multilevel Modelling of the Number of Property Crimes: Household and Area Effects", *Journal of Royal Statistical Society. Series A (Statistics in Society)*, 169, 205-233.

Tseloni, A., and K. Pease (2003), "Repeat Personal Victimization. 'Boosts' or 'Flags'?", *British Journal of Criminology*, 43 196-212.

Tseloni, A., and K. Pease (2004), "Repeat Personal Victimization. Random Effects, Event Dependence and Unexplained Heterogeneity", *British Journal of Criminology*, 44, 931-945.

Turner, C. F., L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck, and F. L. Sonenstein (1998), "Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology", *Science*, 280, 867-873.

Walby S., and J. Allen (2004), "Domestic Violence, Sexual Assault and Stalking: Findings from the British Crime Survey", Home Office Research Study 276.

Wiles, P., J. Simmons, and K. Pease (2003), "Crime Victimization: Its Extent and Communication", *Journal of Royal Statistical Society. Series A (Statistics in Society)*, 166, 247-252.

Winkelmann, R. (2006), "Reforming Health Care: Evidence from Quantile Regressions for Counts", *Journal of Health Economics*, 25, 131-145.

Winkelmann, R. (2008), *Econometric Analysis of Count Data*, 5^{th} ed., Berlin, Springer-Verlag.

Winkelmann, R., and K.F. Zimmermann (1993), "Poisson Logistic Regression", Department of Economics, University of Munich, Working Paper No. 93-18.

Wittebrood, K., and P. Nieubeerta (2000), "Criminal Victimization during One's Life Course: The Effects of Previous Victimization and Patterns on Routine Activities", *Journal of Research in Crime and Delinquency*, 37, 91-122.

Wolfgang, M. (1958), *Patterns in Criminal Homicide*, University of Pennsylvania Press; Philadelphia.

Wooldridge, J. M. (2007), "Inverse Probability Weighted Estimation for General Missing Data Problems," *Journal of Econometrics*, 141, 1281-1301.

Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2^{nd} ed., Cambridge, MIT Press.

Van De Ven, W., and B. Van Praag (1981), "The Demand for Deductibles in Private Health Insurance. A Probit Model with Sample Selection", *Journal of Econometrics*, 17, 229-252.

von Hentig, H. (1940), "Remarks on the Interaction of Perpetrator and Victim ", *Journal of Criminal Law and Criminology*, 31, 303-309.

Table 1. BCS Crime Codes*

	Category	Code	Description	Valid?
0	**Miscellaneous**	01	Refer to Home Office	
		02	Duplicate victim form	
		96	Invalid Victim Form (e.g. no information/no offence)	
1	**Assault**	11	Serious wounding	✓
		12	Other wounding	✓
		13	Common assault	✓
		14	Other assault outside the survey's coverage	
2	**Attempted assault**	21	Attempted assault	✓
3	**Sexual offences**	31	Rape	✓
		32	Serious wounding with sexual motive	✓
		33	Other wounding with sexual motive	✓
		34	Attempted rape	✓
		35	Indecent assault	✓
		39	Sexual offence outside the survey's coverage	
4	**Personal theft**	41	Robbery	✓
		42	Attempted robbery	✓
		43	Snatch theft from the person	✓
		44	Other theft from the person	✓
		45	Attempted theft from the person	✓
		48	Possibly theft but could have been loss/possibly attempted theft, but not certain	
		49	Other robbery or theft from the person outside the survey's coverage	
5	**Burglary/Theft in a dwelling**	50	Attempted burglary to non-connected domestic garage/outhouse	✓
		51	Burglary in a dwelling (nothing taken)	✓
		52	Burglary in a dwelling (Something taken)	✓
		53	Attempted burglary in a dwelling	✓
		54	Possible attempted burglary (insufficient evidence to be sure)	
		55	Theft in a dwelling	✓
		56	Theft from a meter	✓
		57	Burglary from non-connected domestic garage/outhouse – nothing taken	✓
		58	Burglary from non-connected domestic garage/outhouse – something taken	✓
		59	Other burglary, attempted burglary, theft in a dwelling, falling outside the survey's coverage	

* This table is taken by the BCS 2008-09 User Guide pages 19 and 20.

Table 1. Continued

6	**Theft**	60	Theft of car/van	✓
		61	Theft from car/van	✓
		62	Theft of motorbike, motorscooter or moped	✓
		63	Theft from motorbike, motorscooter or moped	✓
		64	Theft of pedal cycle	✓
		65	Theft from outside dwelling (excluding theft of milk bottles)	✓
		66	Theft of milk bottles from outside dwelling	
		67	Other theft	✓
		68	Possible theft, possible lost property	
		69	Other theft/attempted theft falling outside survey's coverage	
7	**Attempted theft**	71	Attempted theft of/from car/van	✓
		72	Attempted theft of/from motorcycle, motorscooter or moped	✓
		73	Other attempted theft	✓
8	**Vandalism**	80	Arson	✓
		81	Criminal damage to a motor vehicle (£20 or under)	✓
		82	Criminal damage to a motor vehicle (over £20)	✓
		83	Criminal damage to the home (£20 or under)	✓
		84	Criminal damage to the home (over £20)	✓
		85	Other criminal damage (£20 or under)	✓
		86	Other criminal damage (over £20)	✓
		87	Possibly criminal/possibly accidental damage/ nuisance with no damage	
		88	Attempted criminal damage (no damage actually achieved)	
		89	Other criminal damage outside survey's coverage	
9	**Threats**	91	Threat to kill/assault made against, but not necessarily to respondent	✓
		92	Sexual threat made against, but not necessarily to respondent	✓
		93	Other threat or intimidation made against, but not necessarily to respondent	✓
		94	Threats against others, made to the respondent	✓
		97	Other threats/intimidation outside survey's coverage	

Table 2. Count Data Tabulations for each Crime Group

	Acquisitive Crime													
	Total		Inside Burglary		Outside Burglary		Vehicle Theft		Inside Theft		Outside Theft		Other Theft	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
0	40,738	*86.87*	45,805	*97.68*	46,421	*98.99*	43,832	*93.47*	46,774	*99.75*	45,663	*97.38*	46,068	*98.24*
1	4,646	*9.91*	921	*1.96*	407	*0.87*	2,496	*5.32*	86	*0.18*	1,001	*2.13*	742	*1.58*
2	950	*2.03*	97	*0.21*	43	*0.09*	398	*0.85*	18	*0.04*	148	*0.32*	55	*0.12*
3	296	*0.63*	26	*0.06*	15	*0.03*	99	*0.21*	4	*0.01*	41	*0.09*	16	*0.03*
4	120	*0.26*	12	*0.03*	1	*0.00*	35	*0.07*	4	*0.01*	19	*0.04*	6	*0.01*
5	47	*0.1*	8	*0.02*	0	*0.00*	19	*0.04*	0	*0.00*	5	*0.01*	3	*0.01*
6	29	*0.06*	4	*0.01*	2	*0.00*	1	*0.00*	3	*0.01*	4	*0.01*	1	*0.00*
7	17	*0.04*	3	*0.01*	1	*0.00*	5	*0.01*	1	*0.00*	3	*0.01*	0	*0.00*
8	8	*0.02*	2	*0.00*	0	*0.00*	1	*0.00*	0	*0.00*	1	*0.00*	1	*0.00*
9	5	*0.01*	1	*0.00*	0	*0.00*	9	*0.00*	1	*0.00*	0	*0.00*	0	*0.00*
10+	37	*0.08*	14	*0.01*	3	*0.00*	6	*0.01*	2	*0.00*	8	*0.02*	1	*0.00*

	Criminal Damage						Personal Theft					
	Total *(+ Other, Arson)*		Home		Vehicle		Total		Mugging		Theft	
	N	%	N	%	N	%	N	%	N	%	N	%
0	43,331	*92.40*	45,733	*97.5*	44,421	*94.73*	46,292	*99.00*	46,630	*99.00*	46,549	*99.30*
1	2,418	*5.16*	776	*1.65*	1,776	*3.79*	554	*1.18*	228	*0.49*	333	*0.71*
2	603	*1.29*	190	*0.41*	409	*0.87*	33	*0.07*	23	*0.05*	10	*0.02*
3	264	*0.56*	79	*0.17*	161	*0.34*	4	*0.01*	2	*0.00*	1	*0.00*
4	105	*0.22*	33	*0.07*	48	*0.10*	2	*0.00*	3	*0.01*	0	*0.00*
5	46	*0.10*	12	*0.03*	32	*0.07*	3	*0.01*	4	*0.01*	0	*0.00*
6	41	*0.09*	24	*0.05*	18	*0.04*	3	*0.01*	2	*0.00*	0	*0.00*
7	9	*0.02*	2	*0.00*	3	*0.01*	1	*0.00*	0	*0.00*	0	*0.00*
8	6	*0.01*	4	*0.01*	2	*0.00*	0	*0.00*	0	*0.00*	0	*0.00*
9	8	*0.02*	1	*0.00*	1	*0.00*	0	*0.00*	0	*0.00*	0	*0.00*
10+	62	*0.01*	39	*0.08*	22	*0.05*	1	*0.00*	1	*0.00*	0	*0.00*

Table 3. Descriptive Statistics

Variables		Mean: All	Mean: Native	Mean: Immigrant	Min	Max	Mis
<u>Personal Crime Variables</u>							
Violence by Strangers (Binary)		0.011	0.011	0.013	0	1	
Violence by Strangers (Count)		0.015 (0.20)	0.015 (0.20)	0.018 (0.20)	0	11	
Violence by Acquaintan. (Binary)		0.010	0.011	0.006	0	1	
Violence by Acquaintan. (Count)		0.027 (0.93)	0.027 (0.85)	0.034 (1.47)	0	97	
Domestic Violence (Binary)		0.005	0.005	0.003	0	1	
Domestic Violence (Count)		0.026 (1.30)	0.028 (1.37)	0.007 (0.17)	0	194	
Mugging (Binary)		0.006	0.005	0.008			
Mugging (Count)		0.009 (0.46)	0.009 (0.48)	0.010 (0.12)	0	97	
Other Personal Theft (Binary)		0.007	0.007	0.011			
Other Personal Theft (Count)		0.008 (0.09)	0.007 (0.09)	0.012 (0.112)	0	3	
<u>Household Crime Variables</u>							
Inside Burglary (Binary)		0.023	0.023	0.030	0	1	
Inside Burglary (Count)		0.046 (1.20)	0.044 (1.17)	0.060 (1.48)	0	100	
Outside Burglary (Binary)		*0.010*	*0.010*	*0.007*	0	1	
Outside Burglary (Count)		0.014 (0.47)	0.015 (0.50)	0.010 (0.15)	0	97	
Vehicle Theft (Binary)		0.065	0.064	0.076	0	1	
Vehicle Theft (Count)		0.085 (0.40)	0.083 (0.40)	0.099 (0.41)	0	20	
Inside, Outside, & Other Thefts (Binary)		0.047	0.047	0.046	0	1	
Inside, Outside, & Other Thefts (Count)		0.073 (1.10)	0.074 (1.12)	0.068 (0.94)	0	98	
Home Criminal Damage (Binary)		0.025	0.025	0.022	0	1	
Home Criminal Damage (Count)		0.075 (1.72)	0.077 (1.74)	0.057 (1.48)	0	97	
Vehicle Criminal Damage (Binary)		0.053	0.052	0.055	0	1	
Vehicle Criminal Damage (Count)		0.092 (1.03)	0.090 (0.96)	0.106 (1.53)	0	97	
<u>Respondent's Characteristics</u>							
Immigrant		0.095			0	1	
Age		50.45 (18.58)	51.01 (18.64)	45.17 (17.16)	16	101	66
Gender (female)	Male	0.454	0.455	0.444	0	1	
Marital Status	Married	0.476	0.470	0.527	0	1	20
	Cohabiting	0.088	0.089	0.074	0	1	
	Single	0.204	0.204	0.209	0	1	
	Widowed	0.115	0.119	0.071	0	1	
	Divorced	0.087	0.090	0.067	0	1	
	Separated	0.030	0.027	0.054	0	1	

Table 3. Continued

Variables		Mean			Min	Max	Mis
		All	Native	Immigrant			
Employment Status	Employed	0.562	0.558	0.605	0	1	
	Unemployed	0.017	0.016	0.023	0	1	
	Inactive Student	0.002	0.002	0.004	0	1	64
	Inactive Retired	0.281	0.292	0.176	0	1	
	Inactive Other	0.117	0.112	0.157	0	1	
Education	None	0.283	0.287	0.250	0	1	
	O-level / GCSE	0.199	0.208	0.112	0	1	
	A-lEvel /Apprent.	0.170	0.176	0.113	0	1	81
	Degree /Diploma	0.304	0.289	0.449	0	1	
	Other	0.043	0.040	0.076	0	1	
Ethnic Group	White	0.933	0.976	0.528	0	1	
	Black	0.018	0.006	0.133	0	1	
	Asian	0.031	0.010	0.233	0	1	7
	Chinese / Other	0.012	0.004	0.086	0	1	
	Mixed	0.006	0.004	0.019	0	1	
<u>Hhd Ref Person's Characteristics</u>							
Immigrant		0.095			0	1	
Age		52.60 (17.13)	53.18 (17.10)	47.08 (16.44)	16	101	105
Gender (female)	Male	0.624	0.624	0.622	0	1	
Marital Status	Married	0.513	0.511	0.538	0	1	
	Cohabiting	0.090	0.092	0.070	0	1	
	Single	0.150	0.147	0.181	0	1	23
	Widowed	0.118	0.123	0.074	0	1	
	Divorced	0.095	0.097	0.076	0	1	
	Separated	0.033	0.030	0.061	0	1	
Employment Status	Employed	0.610	0.603	0.675	0	1	
	Unemployed	0.011	0.011	0.016	0	1	
	Inactive Student	0.009	0.007	0.029	0	1	65
	Inactive Retired	0.280	0.291	0.177	0	1	
	Inactive Other	0.089	0.087	0.102	0	1	
<u>Hhd Caracteristics</u>							
Tenure Type (Renters)	Owners	0.702	0.719	0.543	0	1	127
Condition (Bad)	Indifferent	0.219	0.213	0.284	0	1	
	Good	0.416	0.417	0.405	0	1	2746
	Very Good	0.332	0.339	0.264	0	1	
Relative Condition (Same)	Better	0.085	0.085	0.077	0	1	3059
	Worse	0.062	0.062	0.070	0	1	
Accommodation Type	Detached	0.265	0.273	0.179	0	1	
	Semi Detached	0.332	0.339	0.265	0	1	
	Terrace	0.280	0.277	0.316	0	1	2549
	Flat/ Maisonette	0.119	0.107	0.237	0	1	
	Other	0.005	0.005	0.003			
Location (Other)	Main Road	0.142	0.142	0.139	0	1	
	Side Road	0.536	0.535	0.548	0	1	
Number of Adults		1.898 (1.898)	1.881 (0.809)	2.061 (0.984)	1	10	
Lone Parent		0.051	0.051	0.054	0	1	107

Table 3. Continued

Variables		Mean			Min	Max	Mis
		All	Native	Immigrant			
HOURS AWAY		4.587	4.577	4.682	1	6 (index)	127
YEARS HOME		4.902	4.996	4.015	1	7 (index)	4
YEARS AREA		5.475	5.588	4.401	1	7 (index)	1
NEIGHBOR WATCHING PROGRAM		0.272	0.275	0.242	0	1	10743
INCOME	under £10,000	0.202	0.201	0.205	0	1	10026
	£10,000-£19,999	0.224	0.227	0.199	0	1	
	£20,000-£29,999	0.175	0.177	0.157	0	1	
	£30,000-£39,999	0.135	0.136	0.133	0	1	
	£40,000-£49,999	0.095	0.096	0.089	0	1	
	£50,000 or more	0.153	0.150	0.187	0	1	
	nothing	0.016	0.014	0.030	0	1	
NUMBER OF CARS		1.265 (0.924)	1.284 (0.925)	1.091 (0.894)	0	4(+)	
MOTORCYCLE		0.067	0.070	0.039	0	1	
BICYCLE		0.444	0.452	0.370	0	1	
Area Characteristics							
REGIONS	NORTH EAST	0.066	0.070	0.026	0	1	
	NORTH WEST	0.118	0.122	0.076	0	1	
	YORKSHIRE	0.091	0.095	0.060	0	1	
	EAST MIDLANDS	0.111	0.114	0.088	0	1	
	WEST MIDLANDS	0.100	0.101	0.091	0	1	
	EAST OF ENGLAND	0.130	0.129	0.133	0	1	
	LONDON	0.077	0.055	0.290	0	1	
	SOUTH EAST	0.111	0.110	0.123	0	1	
	SOUTH WEST	0.106	0.109	0.076	0	1	
	WALES	0.091	0.096	0.038	0	1	
URBAN		0.744	0.730	0.880	0	1	
INNER CITY		0.079	0.069	0.167	0	1	
DEPRIVATION INDEX		5.232 (2.824)	5.161 (2.80)	5.911 (2.93)	1	10	
10th percentile		0.109	0.110	0.096			
20th		0.108	0.110	0.083			
30th		0.115	0.118	0.084			
40th		0.110	0.113	0.080			
50th		0.098	0.100	0.073			
60th		0.104	0.104	0.107			
70th		0.098	0.097	0.105			
80th		0.090	0.088	0.115			
90th		0.088	0.084	0.126			
100th percentile		0.082	0.077	0.126			

Standard deviations are presented in parentheses.

Table 4. The Risk of Inside Burglary plus Attempts

INSIDE BURGLARY (including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
HRP IMMIGRANT	0.123***	*0.040*	0.048	*0.043*	0.004	*0.044*	-0.022	*0.046*
DEPRIVATION			0.046***	*0.005*	0.025***	*0.005*	0.027***	*0.006*
LONDON			0.024	*0.048*	0.010	*0.049*	-0.017	*0.052*
URBAN			0.237***	*0.036*	0.203***	*0.036*	0.215***	*0.037*
INNER CITY			-0.010	*0.045*	-0.045	*0.046*	-0.040	*0.047*
HRP AGE					-0.012***	*0.001*	-0.008***	*0.001*
HRP MALE					-0.037	*0.029*	-0.021	*0.030*
HRP MARRIED					-0.111***	*0.030*	-0.101***	*0.034*
HRP EMPLOYED					-0.146***	*0.033*	-0.106***	*0.038*
OWNERS					-0.136***	*0.03*	-0.103***	*0.034*
CONDITION, TYPE LOCATION, NUM ADULTS, LONE PARENT, HOURS UNOCCUPIED YEARS IN HOME/AREA, WATCHING NEIGHBORHOOD, INCOME, EDUCATION							√	
CONSTANT	-2.004***	*0.013*	-2.448***	*0.040*	-1.474***	*0.075*	-1.510***	*0.135*
Log Likelihood	-5,165.03		-5,061,08		-4,897.71		-4,822.63	
R^2	0.0009		0.0193		0.0479		0.0615	
N	46,810		46,810		46,588		46,525	

Table 5. The Risk of Outside Burglary plus Attempts

OUTSIDE BURGLARY (including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
HRP IMMIGRANT	-0.160**	*0.068*	-0.173**	*0.072*	-0.180**	*0.073*	-0.182**	*0.076*
DEPRIVED			0.038***	*0.007*	0.042***	*0.007*	0.043***	*0.008*
LONDON			-0.091	*0.075*	-0.082	*0.076*	-0.123	*0.078*
URBAN			0.092**	*0.044*	0.082**	*0.044*	0.098**	*0.045*
INNER CITY			-0.057	*0.067*	-0.050	*0.067*	-0.045	*0.067*
HRP AGE					-0.005***	*0.001*	-0.003*	*0.002*
HRP MALE					-0.030	*0.040*	-0.006	*0.043*
HRP MARRIED					0.062	*0.040*	0.021	*0.048*
HRP EMPLOYED					-0.008	*0.050*	-0.033	*0.054*
OWNERS					0.139***	*0.046*	0.088*	*0.051*
CONDITION, TYPE LOCATION, NUM ADULTS, LONE PARENT, HOURS UNOCCUPIED YEARS IN HOME/AREA, WATCHING NEIGHBORHOOD, INCOME, EDUCATION							√	
CONSTANT	-2.310***	*0.017*	-2.580***	*0.048*	-2.449***	*0.106*	-2.972***	*0.190*
Log Likelihood	-2,636.27		-2,613.27		-2582.01		-2535.72	
R^2	0.0012		0.0099		0.0157		0.0298	
N	46,810		46,810		46,588		46,525	
Assimilation								
IMMIGRANT	-0.286**	*0.116*	-0.322***	*0.121*	-0.397***	*0.127*	-0.376***	*0.135*
IMMIGRANT'S NUMBER OF YEARS IN COUNTRY	0.003	*0.003*	0.005	*0.003*	0.007**	*0.003*	0.006*	*0.004*
IMMIGRANT (plus HRPAGE)	-0.404***	*0.118*	-0.419***	*0.123*				
IMMIGRANT'S NO. YEARS IN COUNTRY (plus HRPAGE)	0.007***	*0.003*	0.008**	*0.003*				

Robust standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level

Table 6. The Risk of Personal Theft

PERSONAL THEFT (Including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
IMMIGRANT	0.190***	*0.047*	0.113*	*0.059*	0.036	*0.060*	0.030	*0.060*
MALE			-0.082**	*0.032*	-0.082**	*0.033*	-0.076**	*0.035*
AGE 26 – 35			-0.339***	*0.053*	-0.341***	*0.054*	-0.211***	*0.059*
AGE 36 – 45			-0.472***	*0.054*	-0.476***	*0.055*	-0285***	*0.065*
AGE 45 – 56			-0.470***	*0.057*	-0.474***	*0.058*	-0.273***	*0.069*
AGE 56 – PLUS			-0.479***	*0.046*	-0.477***	*0.046*	-0.196***	*0.073*
BLACK			0.186*	*0.097*	0.064	*0.099*	0.012	*0.100*
ASIAN & OTHER			-0.117	*0.083*	-0.159*	*0.084*	-0.163*	*0.087*
MIXED			-0.142	*0.202*	-0.217	*0.202*	-0.389*	*0.228*
DEPRIVATION			0.022***	*0.007*	0.031***	*0.007*	0.027***	*0.008*
URBAN			0.138***	*0.043*	0.076*	*0.045*	0.073	*0.045*
INNER CITY			0.191***	*0.052*	0.129**	*0.053*	0.111**	*0.054*
NORTH EAST					-0.441***	*0.083*	-0.402***	*0.085*
NORTH WEST					-0.373***	*0.068*	-0.319***	*0.069*
YORKSHIRE					-0.364***	*0.073*	-0.292***	*0.074*
EAST MIDLANDS					-0.361***	*0.070*	-0.304***	*0.071*
WEST MIDLANDS					-0.322***	*0.069*	-0.259***	*0.070*
EAST OF ENGLAND					-0.311***	*0.067*	-0.259***	*0.069*
SOUTH EAST					-0.156**	*0.065*	-0.118*	*0.066*
SOUTH WEST					-0.391***	*0.074*	-0.334***	*0.075*
WALES					-0.504***	*0.083*	-0.440***	*0.084*
EDUCATION, MARITAL, EMPLOYMENT, TENURE, INCOME							√	
CONSTANT	-2.254***	*0.017*	-2.083***	*0.062*	-1.760***	*0.082*	-2.062***	*0.117*
Log Likelihood	-3,198.49		-3,090.00		-3577.29		-2,974.69	
R^2	0.0024		0.0362		0.0467		0.0662	
N	46,827		46,820		46,820		46,567	
$\Pr(Y = \widehat{1\|X, Im} = 1)$	0.0195		0.0116		0.0093		0.0040	
$\Pr(Y = \widehat{1\|X, Im} = 0)$	0.0121		0.0086		0.0084		0.0036	
Diff	0.0074***		0.0030*		0.0009		0.0003	
(se)	(0.0021)		(0.0018)		(0.0015)		(0.0007)	
Ratio	1.612		1.355		1.104		1.094	

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 7. The Risk of Violent Victimization

TOTAL ASSAULT (Including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
IMMIGRANT	-0.068	*0.044*	-0.165***	*0.046*	-0.157***	*0.049*	-0.151***	*0.051*
MALE			0.179***	*0.026*	0.185***	*0.026*	0.240***	*0.029*
AGE 20 – 24			-0.195***	*0.056*	-0.216***	*0.057*	-0.200***	*0.063*
AGE 25 – 34			-0.426***	*0.049*	-0.432***	*0.049*	-0.354***	*0.061*
AGE 35 – 44			-0.671***	*0.049*	-0.663***	*0.049*	-0.542***	*0.064*
AGE 45 – 54			-0.824***	*0.053*	-0.813***	*0.053*	-0.659***	*0.070*
AGE 55 – 64			-1.119***	*0.059*	-1.107***	*0.06*	-0.924***	*0.079*
AGE 65 – 74			-1.361***	*0.074*	-1.353***	*0.075*	-1.139***	*0.098*
AGE 75 – PLUS			-1.897***	*0.138*	-1.889***	*0.139*	-1.757***	*0.168*
DEPRIVATION					0.025***	*0.005*	0.010*	*0.006*
URBAN					0.059***	*0.034*	0.066*	*0.035*
INNER CITY					0.023	*0.048*	0.010	*0.049*
REGIONS					√		√	
EDUCATION, MARITAL, EMPLOYMENT, TENURE, INCOME, LONE PARENT, HHD MEMBERS							√	
CONSTANT	-1.948***	*0.013*	-1.306***	*0.043*	-1.654***	*0.078*	-1.861***	*0.127*
Log Likelihood	-5,536.52		-4,989.29		-4,959.87		-4,777.29	
R^2	0.0002		0.1002		0.1055		0.1275	
N	46,827		46,827		46,827		46,532	
$\Pr(Y = \widehat{1\|X}, Im = 1)$	0.0219		0.0248		0.0228		0.0265	
$\Pr(Y = \widehat{1\|X}, Im = 0)$	0.0258		0.0361		0.0327		0.0372	
Diff	-0.0038*		-0.0113***		-0.0100***		-0.0107***	
	(0.0023)		(0.0028)		(0.0029)		(0.0036)	
Ratio	0.8512		0.6872		0.6971		0.7121	

Table 8. The Risk of Violent Victimization – Including Ethnic Group Dummies

TOTAL ASSAULT (Including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
IMMIGRANT	-0.093*	*0.054*	-0.114**	*0.053*	-0.103**	*0.055*	-0.093*	*0.057*
BLACK	0.005	*0.102*	-0.130	*0.105*	-0.149	*0.108*	-0.216**	*0.110*
ASIAN	0.053	*0.081*	-0.156**	*0.080*	-0.174**	*0.081*	-0.170**	*0.087*
CHINESE OR OTHER	0.108	*0.115*	-0.022	*0.123*	-0.033	*0.124*	-0.037	*0.126*
MIXED	0.242*	*0.139*	-0.020	*0.145*	-0.035	*0.146*	-0.116	*0.153*
Log Likelihood	-5,534.43		-4,980.17		-4,950.12		-4,769.33	
R^2	0.0006		0.1006		0.1061		0.1283	
N	46,820		46,818		46,818		46,526	
$\Pr(Y = \widehat{1\|X}, Im = 1)$	0.0205		0.0281		0.0262		0.0304	
$\Pr(Y = \widehat{1\|X}, Im = 0)$	0.0256		0.0363		0.0331		0.0373	
Diff	-0.0051*		-0.0082***		-0.0069**		-0.0070*	
	(0.0023)		(0.0035)		(0.0034)		(0.0041)	
Ratio	0.8015		0.7740		0.7919		0.8132	

For both Table 7 and Table 8, Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 9. The Risk of Domestic Violence

DOMESTIC (Including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
Immigrant	-0.177*	*0.091*	-0.248**	*0.100*	-0.219**	*0.107*	-0.207*	*0.112*
MALE			-0.438***	*0.056*	-0.419***	*0.057*	-0.337***	*0.061*
AGE 26 – 35			0.123*	*0.069*	0.167**	*0.074*	0.114	*0.081*
AGE 36 – 45			-0.131*	*0.073*	-0.121	*0.082*	-0.138	*0.09*
AGE 45 – 56			-0.346***	*0.087*	-0.335***	*0.104*	-0.283**	*0.115*
AGE 56 – PLUS			-0.794***	*0.090*	-0.682***	*0.115*	-0.646***	*0.14*
DEPRIVATION			0.027***	*0.010*	0.017	*0.010*	-0.003	*0.011*
URBAN			-0.005	*0.060*	-0.015	*0.062*	0.003	*0.063*
INNER CITY			0.086	*0.085*	0.090	*0.087*	0.065	*0.089*
REGIONS			√		√		√	
COHABITING					0.129	*0.097*	0.136	*0.098*
SINGLE					0.363***	*0.073*	0.238***	*0.084*
WIDOWED					-0.093	*0.195*	-0.204	*0.194*
DIVORCED					0.621***	*0.084*	0.450***	*0.096*
SEPARATED					0.882***	*0.092*	0.711***	*0.106*
EDUCATION, EMPLOYMENT, TENURE, INCOME, LONE PARENT, HHD MEMBERS							√	
CONSTANT	-2.556***	*0.023*	-2.668***	*0.150*	-2.954***	*0.166*	-2.673***	*0.226*
Log Likelihood	-1,492.38		-1,345.71		-1,283.47		-1,232.80	
R^2	0.0014		0.0996		0.1412		0.1684	
N	46,827		46,827		46,811		46,532	
$\Pr(Y = \widehat{1\|X, Im} = 1)$	0.0031		0.0041		0.0017		0.0011	
$\Pr(Y = \widehat{1\|X, Im} = 0)$	0.0053		0.0084		0.0034		0.0022	
Diff	-0.0022**		-0.0042**		-0.0017**		-0.0011*	
(se)	(0.001)		(0.002)		(0.001)		(0.001)	
ratio	0.593		0.494		0.505		0.511	

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 10. The Risk of Victimization suffered by Acqaintances

By Acquaintances (Including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
Immigrant	-0.209***	*0.070*	-0.299***	*0.078*	-0.274***	*0.081*	-0.165*	*0.087*
Male			0.198***	*0.037*	0.219***	*0.040*	0.222***	*0.041*
Age 26 – 35			-0.384***	*0.050*	-0.225***	*0.056*	-0.224***	*0.056*
Age 36 – 45			-0.589***	*0.053*	-0.391***	*0.064*	-0.398***	*0.064*
Age 45 – 56			-0.722***	*0.061*	-0.484***	*0.077*	-0.494***	*0.077*
Age 56 – plus			-1.18***	*0.064*	-0.896***	*0.095*	-0.91***	*0.096*
Deprivation			0.031***	*0.008*	0.011	*0.008*	0.013	*0.008*
Urban			-0.022	*0.047*	-0.012	*0.049*	-0.005	*0.049*
Inner City			0.009	*0.068*	-0.003	*0.069*	-0.003	*0.070*
Regions			√		√		√	
Education, Marital, Employment, Tenure, Income, Lone Parent, Hhd members					√		√	
Black							-0.193	*0.162*
Asian							-0.582***	*0.184*
Other							-0.188	*0.200*
Mixed							-0.056	*0.189*
Constant	-2.300***	*0.018*	-2.121***	*0.105*	-2.348***	*0.161*	-2.332***	*0.163*
Log Likelihood	-2,675.70		-2,392.23		-2,297.46		-2,289.76	
R^2	0.0019		0.1076		0.1301		0.1330	
N	46,827		46,827		46,532		46,526	
$\Pr\,(Y = \widehat{1\|X}, Im = 1)$	0.0060		0.0049		0.0032		0.0046	
$\Pr\,(Y = \widehat{1\|X}, Im = 0)$	0.0107		0.0112		0.0072		0.0074	
Diff	-0.0047*** (0.0013)		-0.0063*** (0.0016)		-0.0039*** (0.0013)		-0.0028** (0.0014)	
Ratio	0.564		0.438		0.451		0.625	

Table 11. The Risk of Victimization suffered by Strangers

By Strangers (including Attempts)	1		2		3		4	
Probit	coeff	se	coeff	se	coeff	se	coeff	se
Immigrant	0.084	*0.053*	0.010	*0.056*	-0.004	*0.059*	-0.009	*0.061*
Male			0.441***	*0.038*	0.444***	*0.038*	0.411***	*0.041*
Age 26 – 35			-0.341***	*0.05*	-0.339***	*0.051*	-0.279***	*0.057*
Age 36 – 45			-0.518***	*0.052*	-0.506***	*0.052*	-0.402***	*0.062*
Age 45 – 56			-0.672***	*0.06*	-0.658***	*0.061*	-0.543***	*0.074*
Age 56 – plus			-1.029***	*0.058*	-1.012***	*0.058*	-0.795***	*0.088*
Deprivation					0.011	*0.007*	0.015*	*0.008*
Urban					0.138***	*0.048*	0.136***	*0.049*
Inner City					-0.015	*0.066*	-0.007	*0.067*
Regions					√		√	
Other regressors (as for by acquaintance crime)							√	
Constant	-2.304***	*0.018*	-2.013***	*0.041*	-2.223***	*0.094*	-2.565***	*0.154*
Log Likelihood	-2,806.62		-2,534.25		-2,525.73		-2,456.38	
R^2	0.0004		0.0974		0.1005		0.1127	

For both Table 10 and Table 11, Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 12. A Trivariate Probit Model for Violent Victimization

Trivariate Probit (300 draws)	1st Equation Domestic		2nd Equation By Acquaintances		3rd Equation By Strangers	
	coeff	se	coeff	se	coeff	se
IMMIGRANT	-0.213**	*0.089*	-0.298***	*0.077*	-0.013	*0.059*
MALE	-0.433***	*0.054*	0.195***	*0.037*	0.445***	*0.038*
AGE 26 – 35	0.118*	*0.068*	-0.382***	*0.050*	-0.338***	*0.051*
AGE 36 – 45	-0.142**	*0.072*	-0.586***	*0.053*	-0.504***	*0.052*
AGE 45 – 56	-0.324***	*0.081*	-0.724***	*0.061*	-0.665***	*0.061*
AGE 56 – PLUS	-0.805***	*0.091*	-1.178***	*0.063*	-1.008***	*0.058*
DEPRIVATION	0.027***	*0.010*	0.030***	*0.008*	0.011	*0.007*
URBAN	0.002	*0.060*	-0.023	*0.047*	0.140***	*0.048*
INNER CITY	0.086	*0.084*	0.004	*0.068*	-0.010	*0.066*
NORTH EAST	0.230	*0.152*	0.271***	*0.101*	0.088	*0.093*
NORTH WEST	0.255*	*0.135*	0.107	*0.097*	0.019	*0.083*
YORKSHIRE	0.421***	*0.133*	0.177*	*0.099*	-0.023	*0.090*
EAST MIDLANDS	0.455***	*0.131*	0.166**	*0.098*	0.112	*0.082*
WEST MIDLANDS	0.344***	*0.134*	0.237**	*0.096*	0.021	*0.085*
EAST OF ENGLAND	0.211*	*0.136*	0.086	*0.099*	0.027	*0.083*
SOUTH EAST	0.306**	*0.136*	0.206**	*0.099*	0.051	*0.085*
SOUTH WEST	0.404***	*0.136*	0.069	*0.104*	0.046	*0.087*
WALES	0.391***	*0.139*	0.183*	*0.103*	0.033	*0.091*
CONSTANT	-2.660***	*0.144*	-2.118***	*0.105*	-2.231***	*0.094*
Log Likelihood	-6,254.93					
N	46,827					
$\hat{\rho}_{12}$	0.153***	*0.058*	LR test for ρ_{12} =ρ_{13}=ρ_{23}=0			
$\hat{\rho}_{13}$	0.013	*0.059*	χ^2 =17.48			
$\hat{\rho}_{13}$	0.142***	*0.046*	P-value=0.0006			

Quasi-Maximum Likelihood standard errors are presented in *italics*
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 13. Comparisons Between Face-to-Face and Self-Reports

SELF-COMPLETION DOMESTIC CRIME	Face-to-face Simple Probit (16 – 59)		No Sample Selection Correction		Correcting for Sample Selection (given acceptance)		Correcting for Sample Selection (16 - 59)	
Crime Equation								
	Coeff	se	Coeff	se	Coeff	se	Coeff	se
IMMIGRANT	-0.284***	*0.103*	-0.223***	*0.062*	-0.214***	*0.074*	-0.258***	*0.066*
MALE	-0.434***	*0.057*	-0.191***	*0.032*	-0.190***	*0.032*	-0.192***	*0.031*
DEPRIVATION	0.028***	*0.010*	0.035***	*0.006*	0.037***	*0.008*	0.032***	*0.007*
URBAN	0.011	*0.063*	0.009	*0.039*	0.008	*0.039*	0.009	*0.039*
INNER CITY	0.061	*0.088*	0.028	*0.057*	0.028	*0.057*	0.026	*0.057*
AGE & REGIONAL DUMMIES	√		√		√		√	
CONSTANT	-2.667***	*0.152*	-1.824***	*0.086*	-1.820***	*0.089*	-1.844***	0.086
Selection Equation								
IMMIGRANT					-0.235***	*0.031*	-0.440***	*0.026*
MALE					-0.048**	*0.019*	-0.054***	*0.017*
DEPRIVATION					-0.050***	*0.004*	-0.029***	*0.004*
URBAN					0.013	*0.024*	0.016	*0.022*
INNER CITY					0.009	*0.036*	0.004	*0.032*
AGE & REGIONAL DUMMIES					√		√	
LANGUAGE DIFFICULTIES					-0.877***	*0.047*		
OTHER PRESENT							-0.159***	*0.019*
NO QUALIFICATION							-0.632***	*0.022*
CONSTANT					1.818***	*0.055*	1.451***	*0.048*
$\hat{\rho}$ (P-value from Wald Test)					-0.077	(0.816)	0.233	(0.215)
Log Likelihood	-1,254.06		-3,650.27		-14,448.08		-17,519.96	
N Total	30,711		24,363		28,339		30,324	
N Uncensored					24,344		24,346	
N Censored					3,995		5,978	
$\Pr(\widehat{Y = 1 \mid X, Im = 1})$	0.0040		0.0236		0.0252		0.0197	
$\Pr(\widehat{Y = 1 \mid X, Im = 0})$	0.0090		0.0392		0.0406		0.0358	
Diff	-0.0050***		-0.0155***		-0.0155***		-0.0161***	
(se)	(0.0018)		(0.0039)		(0.004)		(0.0035)	
Ratio	0.447		0.604		0.619		0.550	

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 14. The Presence of the Partner – Probit Estimates[⊕]

	Coeff	se	Coeff	se	Coeff	se
Domestic Face-to-face	1		2		3	
IMMIGRANT	-0.227**	*0.102*	-0.281**	*0.110*	-0.275**	*0.119*
PARTNER PRESENT	-0.408***	*0.093*	-0.314***	*0.100*	-0.125	*0.113*
IMMIGRANT & PARTNER PRESENT	0.446**	*0.235*	0.335	*0.254*	0.377	*0.256*
Domestic Self-completion	1		2		3	
IMMIGRANT	-0.187***	*0.062*	-0.209***	*0.066*	-0.186***	*0.070*
PARTNER PRESENT	-0.157***	*0.054*	-0.118**	*0.056*	0.067	*0.061*
IMMIGRANT & PARTNER PRESENT	0.006	*0.183*	-0.067	*0.186*	-0.023	*0.190*
Acquaintance	1		2		3	
IMMIGRANT	-0.204***	*0.075*	-0.285***	*0.083*	-0.249***	*0.084*
PARTNER PRESENT	-0.237***	*0.058*	-0.131**	*0.062*	0.005	*0.068*
IMMIGRANT & PARTNER PRESENT	0.007	*0.213*	-0.049	*0.226*	-0.021	*0.226*

Table 15. Network Effects and Assimilation Patterns

	Linear Trend (Acquaintnaces) (1)		Quadratic Trend (Acquaintances) (2)		Linear Trend (Domestic) (4)		Quadratic Trend (Strangers NO RMC) (5)	
Probit	Coeff	se	Coeff	se	Coeff	se	Coeff	se
IMMIGRANT	-0.405***	*0.106*	-0.646***	*0.165*	-0.587***	*0.157*	-0.342**	*0.151*
NUMBER OF YEARS IN COUNTRY	0.006*	*0.004*	0.038**	*0.017*	0.015***	*0.005*	0.029*	*0.015*
NUMBER OF YEARS IN COUNTRY^2			-0.0006*	*0.0004*			-0.0006**	*0.0003*
MALE	0.198***	*0.037*	0.199***	*0.037*	-0.439***	*0.055*	0.439***	*0.039*
AGE 26 – 35	-0.385***	*0.050*	-0.387***	*0.050*	0.124*	*0.069*	-0.343***	*0.052*
AGE 36 – 45	-0.594***	*0.053*	-0.598***	*0.053*	-0.142*	*0.073*	-0.537***	*0.055*
AGE 45 – 56	-0.730***	*0.061*	-0.732***	*0.061*	-0.371***	*0.087*	-0.647***	*0.062*
AGE 56 – PLUS	-1.191***	*0.065*	-1.184***	*0.064*	-0.833***	*0.092*	-1.011***	*0.060*
DEPRIVATION	0.031***	*0.007*	0.031***	*0.007*	0.027***	*0.010*	0.005	*0.007*
URBAN	-0.022	*0.047*	-0.022	*0.047*	-0.005	*0.060*	0.142***	*0.049*
REGIONS	√		√		√		√	
CONSTANT	-2.116***	*0.105*	-2.121***	*0.105*	-2.649***	*0.149*	-2.179***	*0.096*
Log Likelihood	-2,391.29		-2,389.52		-1,341.63		-2,373.892	
R^2	0.1079		0.1086		0.1022		0.1004	
N	46,808		46,808		46,808		46,771	
Marginal Effects[ᵛ]								
1) $\Pr\,(\widehat{Y = 1 \mid X, Im = 0})$	0.0190		0.0190		0.0170		0.0173	
2) $\Pr\,(\widehat{Y = 1 \mid X, Im = 1})$	0.0066		0.0033		0.0034		0.0071	
Difference (at 0 years) 2 - 1	-0.0124***	*0.0028*	-0.0157***	*0.0030*	-0.0136***	*0.0033*	-0.0103***	*0.0033*
Difference (after 10 years)	-0.0112***	*0.0026*	-0.0109***	*0.0027*	-0.0117***	*0.0031*	-0.0040	*0.0027*
Difference (after 20 years)	-0.0098***	*0.0025*	-0.0055	*0.0038*	-0.0088***	*0.0030*	0.0008	*0.0041*
Difference (after 30 years)	-0.0082***	*0.0028*	-0.0029	*0.0045*	-0.0047	*0.0034*	0.0013	*0.0049*
Difference (after 40 years)	-0.0064*	*0.0035*	-0.0051	*0.0050*	0.0012	*0.0052*	-0.0029	*0.0044*
Difference (after 50 years)	-0.0043	*0.0049*	-0.0104	*0.0063*	0.0092	*0.0090*	-0.0092**	*0.0045*
Difference (after 60 years)	-0.0019	*0.0068*	-0.0154***	*0.0056*	0.0201	*0.0152*	-0.0142***	*0.0039*
Difference (after 70 years)	0.0009	*0.0093*	-0.0181***	*0.0037*	0.0344	*0.0240*	-0.0166***	*0.0030*

For both Tables 14 and 15, Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***), (**), (*) denote statistical significance at 1%, 5%, 10% significance level, respectively.

[⊕] Specification 2 also includes age, gender, and area dummies. Specification 3 further includes marital, education, and

[ᵛ] The marginal effects for crime by acquaintances and crime by strangers are calculated for a male, 26 to 35 years old, in an average deprived and urban area in East of England. Note that the average 'number of years in the country' for an immigrant is 26 years. For domestic crime the marginal effects are calculated for a person with characteristics as before, but female.

Table 16. Tabulation of Racially Motivated Crime by Relationship Type

	Racially Motivated Crime			
	No		*Yes*	
DOMESTIC	237	(99.58%)	1	(0.42%)
BY ACQUAINTANCES	481	(100.0%)	0	(0.00%)
BY STRANGERS	472	(92.73%)	37	(7.27%)
Immigrants	*42*	(71.19%)	*17*	(28.81%)
Natives	*430*	(95.56%)	*20*	(4.44%)

Table 17. Mean Comparison of Racially Motivated Crime by Relationship Type

Mean Comparisons	Immigrants	Natives	Diff	Ratio
CRIME BY STRANGERS				
Without controlling for RMC	0.0132	0.0106	0.0026	1.244
After controlling for RMC	0.0094	0.0102	-0.0007	0.930
BY ACQUAINTANCES	0.0060	0.0107	-0.0047	0.564***
DOMESTIC	0.0031	0.0053	-0.0022	0.593*

Table 18. Probit Models before and after controlling for Racially Motivated Crime

	Strangers (No Control for RMC)		By Strangers (Control for RMC)		By Acquaintances		Domestic	
Probit	Coeff	se	Coeff	se	Coeff	se	Coeff	se
IMMIGRANT	-0.004	*0.059*	-0.122*	*0.067*	-0.299***	*0.078*	-0.248**	*0.100*
MALE	0.444***	*0.038*	0.438***	*0.039*	0.198***	*0.037*	-0.438***	*0.056*
AGE 26 – 35	-0.339***	*0.051*	-0.341***	*0.052*	-0.384***	*0.05*	0.123*	*0.069*
AGE 36 – 45	-0.506***	*0.052*	-0.531***	*0.054*	-0.589***	*0.053*	-0.131*	*0.073*
AGE 45 – 56	-0.658***	*0.061*	-0.641***	*0.062*	-0.722***	*0.061*	-0.346***	*0.087*
AGE 56 – PLUS	-1.012***	*0.058*	-1.013***	*0.06*	-1.180***	*0.064*	-0.794***	*0.09*
DEPRIVATION	0.011	*0.007*	0.005	*0.007*	0.031***	*0.008*	0.027***	*0.01*
URBAN	0.138***	*0.048*	0.141***	*0.049*	-0.022	*0.047*	-0.005	*0.06*
INNER CITY	-0.015	*0.066*	-0.013	*0.069*	0.009	*0.068*	0.086	*0.085*
REGIONS◊	√		√		√		√	
CONSTANT	-2.223***	*0.094*	-2.176***	*0.095*	-2.121***	*0.105*	-2.668***	*0.150*
Log Likelihood	-2,525.73		-2,375.9717		-2,392.23		-1,345.71	
R^2	0.1005		0.0997		0.1076		0.0996	
N	46,827		46,827		46,827		46,827	
$\Pr(Y = \widehat{1\mid X}, Im = 1)$	0.0189		0.0129		0.0049		0.0041	
$\Pr(Y = \widehat{1\mid X}, Im = 0)$	0.0191		0.0175		0.0112		0.0084	
Diff	-0.0002		-0.0046**		-0.0063***		-0.0042**	
(se)	(0.0027)		(0.0023)		(0.0016)		(0.0020)	
Ratio	0.990		0.736		0.438		0.494	

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

◊ Regions' effect is jointly insignificant for crime by strangers.

Table 19. Interaction Terms between Immigration and Ethnic Status

Immigration & Ethnic Background	Total Assaults		Domestic		Acquaintances		Strangers (No RMC)	
	Coeff	se	Coeff	se	Coeff	se	Coeff	se
IMMIGRANT	-0.110*	*0.063*	-0.463**	*0.193*	-0.213**	*0.097*	-0.079	*0.153*
WHITE			-0.063	*0.130*			0.205*	*0.122*
BLACK	-0.184	*0.173*			0.050	*0.198*		
ASIAN	-0.236*	*0.122*			-0.670***	*0.255*		
CHINESE & OTHER	0.277	*0.173*			-0.344	*0.37*		
MIXED	-0.065	*0.173*			-0.048	*0.213*		
WHITE & IMMIGRANT			0.331	*0.232*			0.056	*0.174*
BLACK & IMMIGRANT	0.064	*0.220*			-0.427	*0.321*		
ASIAN & IMMIGRANT	0.107	*0.165*			0.274	*0.335*		
(CHINESE & OTHER) & IMMIGRANT	-0.535**	*0.244*			0.260	*0.443*		
MIXED & IMMIGRANT	0.166	*0.323*			0.482	*0.388*		
Log-Likelihood	-4,978.06		-1,344.55		-2,382.44		-2,363.39	
R^2	0.1010		0.1003		0.1112		0.1013	

Table 20. Interaction Terms between Immigration Status and Regional Dummies

Immigration & Location	***Regions***							
	Total Assault		Domestic		Acquaintances		Strangers (No RMC)	
	Coeff	se	Coeff	se	Coeff	se	Coeff	se
IMMIGRANT	-0.324***	*0.115*	-0.271	*0.269*	-0.521***	*0.192*	-0.206	*0.140*
NORTH	0.098	*0.066*	0.312**	*0.146*	0.121	*0.094*	-0.046	*0.084*
MIDLANDS	0.123*	*0.065*	0.372**	*0.146*	0.121	*0.094*	-0.010	*0.082*
WALES	0.138*	*0.076*	0.378**	*0.161*	0.128	*0.108*	-0.005	*0.099*
SOUTH	0.112	*0.068*	0.346**	*0.15*	0.108	*0.099*	-0.015	*0.086*
IMMIGRANT & NORTH	0.278*	*0.149*	0.157	*0.330*	0.332	*0.240*	0.129	*0.203*
IMMIGRANT & MIDLANDS	0.154	*0.140*	-0.115	*0.322*	0.225	*0.232*	0.134	*0.176*
IMMIGRANT & WALES	0.171	*0.252*	0.205	*0.458*	0.491	*0.351*	-0.104	*0.399*
IMMIGRANTS & SOUTH	0.153	*0.158*	0.066	*0.343*	0.177	*0.267*	0.065	*0.207*
DEPRIVATION INDEX								
DEPRIVATION*IMMIGRANT								
Log-Likelihood	-4,993.37		-1,349.82		-2.395.99		-2.378.46	
R^2	0.0995		0.0968		0.1062		0.0988	

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 21. The effect of being an Immigrant on perceived Seriousness

SERIOUSNESS	1		2		3		4		5	
Ordinal Probit	Coeff	se	Coeff	se	Coeff	se	Coeff	se	Coeff	se
IMMIGRANT	0.270***	*0.037*	0.243***	*0.039*	0.281***	*0.040*	0.063	*0.043*	0.003	*0.051*
BLACK							0.515***	*0.085*	0.396***	*0.138*
ASIAN							0.481***	*0.063*	0.371***	*0.096*
OTHER							0.255**	*0.105*	0.109	*0.172*
MIXED							-0.123	*0.132*	-0.099	*0.145*
IMMIGRANT & BLACK									0.231	*0.175*
IMMIGRANT & ASIAN									0.224*	*0.128*
IMMIGRANT & OTHER									0.278	*0.218*
IMMIGRANT & MIXED									-0.075	*0.341*
AGE DUMMIES, MALE, DEPRIVATION, URBAN, INNER CITY, REGIONS			√		√		√		√	
MARITAL, EDUCATION, EMPLOYMENT, INCOME, TENURE					√					
Cutpoint 1	0.443	0.013	0.655	0.063	0.340	0.090	0.689	0.063	0.675	0.064
Cutpoint 2	1.404	0.018	1.629	0.064	1.330	0.091	1.669	0.064	1.656	0.065
Cutpoint 3	2.131	0.029	2.365	0.069	2.082	0.095	2.414	0.069	2.402	0.07
Log Likelihood	-9,774.44		-9,675.22		-9,495.70		-9,626.40		-9,623.63	
N	11,208		11,208		11,148		11,205		11,205	

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level.

Table 22. Distribution of Violent Crime

	Total Violence		Violence Zero Truncated	Domestic		Domestic Zero Truncated	By Acquaintance		Acquaintance Zero Truncated	By Stranger		By Stranger Zero Truncated
Observations	46827		1190	46827		238	46827		481	46827		509
Mean	0.0692		2.7218	0.0264		5.2017	0.0274		2.6632	0.0153		1.4106
Std. Deviation	1.613		9.759	1.3048		17.589	0.9262		8.7547	0.1956		1.2464
Variance	2.6018		95.239	1.7026		309.37	0.8578		76.645	0.0382		1.5535
Skewness	71.044		11.803	103.48		7.4549	85.974		9.0086	26.721		4.7561
Percentiles 75%	0		2	0		3	0		2	0		1
90%	0		4	0		6	0		3	0		2
95%	0		6	0		12	0		5	0		3
99%	1		40	0		97	1		50	1		8
	N	%	%	N	%	%	N	%	%	N	%	%
0	45,637	97.46	-	46,589	99.49	-	46,346	98.97	-	46,318	98.91	-
1	842	1.8	70.76	126	0.27	52.94	349	0.75	72.56	412	0.88	80.94
2	164	0.35	13.78	42	0.09	17.65	67	0.14	13.93	60	0.13	11.79
3	64	0.14	5.38	22	0.05	9.24	24	0.05	4.99	15	0.03	2.95
4	29	0.06	2.44	12	0.03	5.04	7	0.01	1.46	4	0.01	0.79
5	21	0.04	1.76	7	0.01	2.94	10	0.02	2.08	6	0.01	1.18
6	23	0.05	1.93	7	0.01	2.94	6	0.01	1.25	6	0.01	1.18
7	2	0.00	0.17	1	0.00	0.42	1	0.00	0.21	0	0.00	0.00
8	5	0.01	0.42	2	0.00	0.84	1	0.00	0.21	1	0.00	0.20
9	1	0.00	0.08	0	0.00	0.00	1	0.00	0.21	0	0.00	0.00
10	13	0.03	1.09	6	0.01	2.52	3	0.01	0.62	4	0.01	0.79
11	1	0.00	0.08	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
12	2	0.00	0.17	2	0.00	0.84	1	0.00	0.21	1	0.00	0.20
13	1	0.00	0.08	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
15	1	0.00	0.08	1	0.00	0.42	0	0.00	0.00	0	0.00	0.00
20	6	0.01	0.5	2	0.00	0.84	4	0.01	0.83	0	0.00	0.00
24	1	0.00	0.08	1	0.00	0.42	0	0.00	0.00	0	0.00	0.00
25	0	0.00	0.00	0	0.00	0.00	1	0.00	0.21	0	0.00	0.00
26	1	0.00	0.08	0	0.00	0.00	0	0.00	0.00	0	0.00	0.00
40	2	0.00	0.17	1	0.00	0.42	1	0.00	0.21	0	0.00	0.00
48	1	0.00	0.08	1	0.00	0.42	0	0.00	0.00	0	0.00	0.00
50	1	0.00	0.08	0	0.00	0.00	1	0.00	0.21	0	0.00	0.00
60	1	0.00	0.08	0	0.00	0.00	1	0.00	0.21	0	0.00	0.00
75	1	0.00	0.08	1	0.00	0.42	0	0.00	0.00	0	0.00	0.00
97	5	0.01	0.42	2	0.00	0.84	3	0.01	0.62	0	0.00	0.00
100	1	0.00	0.08	1	0.00	0.42	0	0.00	0.00	0	0.00	0.00
194	1	0.00	0.08	1	0.00	0.42	0	0.00	0.00	0	0.00	0.00

Table 23. Poisson, Negative Binomial 2, Censored Poisson

	Logit	Poisson	NegBin2	Censored Poisson				
				5	10	15	20	25
IMMIGRANT	-0.371***	-0.240	-0.360	-0.359***	-0.365**	-0.381**	-0.388**	-0.383**
	0.113	*0.427*	*0.283*	*0.134*	*0.149*	*0.162*	*0.177*	*0.194*
MALE	0.440***	-0.382**	-0.314**	0.236***	0.167	0.106	0.055	0.023
	0.060	*0.183*	*0.136*	*0.070*	*0.082*	*0.090*	*0.099*	*0.106*
AGE 26 – 35	-0.638***	-0.549**	-0.631***	-0.660***	-0.648***	-0.653***	-0.669***	-0.670***
	0.078	*0.237*	*0.171*	*0.093*	*0.109*	*0.119*	*0.129*	*0.136*
AGE 36 – 45	-1.185***	-0.672**	-0.795***	-1.108***	-1.096***	-1.079***	-1.062***	-1.032***
	0.085	*0.283*	*0.207*	*0.101*	*0.117*	*0.131*	*0.147*	*0.157*
AGE 45 – 56	-1.605***	-1.702***	-1.773***	-1.709***	-1.772***	-1.802***	-1.829***	-1.829***
	0.103	*0.356*	*0.217*	*0.118*	*0.133*	*0.145*	*0.159*	*0.172*
AGE 56 – PLUS	-2.810***	-3.244***	-3.285***	-2.976***	-3.024***	-3.037***	-3.060***	-3.062***
	0.112	*0.229*	*0.217*	*0.128*	*0.152*	*0.172*	*0.191*	*0.206*
DEPRIVATION	0.054***	0.042	0.048*	0.067***	0.071***	0.071***	0.070***	0.071***
	0.012	*0.037*	*0.028*	*0.015*	*0.017*	*0.019*	*0.021*	*0.023*
URBAN	0.127	-0.170	-0.365*	0.009	-0.013	-0.037	-0.056	-0.052
	0.079	*0.302*	*0.217*	*0.093*	*0.110*	*0.124*	*0.139*	*0.148*
INNER CITY	0.034	-0.022	0.126	0.035	0.057	0.060	0.072	0.047
	0.106	*0.211*	*0.177*	*0.127*	*0.149*	*0.160*	*0.174*	*0.177*
NORTH EAST	0.530***	0.456	0.306	0.333*	0.361*	0.385*	0.412*	0.412*
	0.163	*0.280*	*0.244*	*0.187*	*0.209*	*0.221*	*0.236*	*0.238*
NORTH WEST	0.265*	0.270	0.205	0.264	0.267	0.260	0.257	0.255
	0.153	*0.230*	*0.212*	*0.177*	*0.189*	*0.191*	*0.192*	*0.194*
YORKSHIRE	0.394**	0.746***	0.468**	0.520***	0.595***	0.638***	0.673***	0.688***
	0.157	*0.272*	*0.239*	*0.182*	*0.199*	*0.209*	*0.220*	*0.224*
EAST MIDLANDS	0.535***	0.858***	0.666***	0.550***	0.579***	0.607***	0.636***	0.667***
	0.151	*0.322*	*0.250*	*0.176*	*0.190*	*0.196*	*0.204*	*0.212*
WEST MIDLANDS	0.447***	1.054***	0.888***	0.569***	0.696***	0.781***	0.840***	0.866***
	0.152	*0.303*	*0.273*	*0.176*	*0.193*	*0.203*	*0.212*	*0.219*
EAST OF ENGLAND	0.230	0.882**	0.664**	0.239	0.337*	0.394**	0.449**	0.465**
	0.155	*0.424*	*0.337*	*0.181*	*0.202*	*0.211*	*0.223*	*0.227*
SOUTH EAST	0.396**	0.730**	0.927**	0.469***	0.482***	0.509***	0.527***	0.545***
	0.156	*0.356*	*0.377*	*0.177*	*0.190*	*0.197*	*0.204*	*0.212*
SOUTH WEST	0.342*	0.764**	0.702**	0.351*	0.464***	0.530***	0.590**	0.644**
	0.160	*0.347*	*0.355*	*0.188*	*0.213*	*0.227*	*0.243*	*0.257*
WALES	0.456***	1.549***	1.178***	0.498***	0.605***	0.701***	0.791***	0.876***
	0.162	*0.477*	*0.357*	*0.191*	*0.220*	*0.241*	*0.264*	*0.284*
CONSTANT	-3.342***	-2.277***	-1.980***	-2.801***	-2.714***	-2.659***	-2.605***	-2.600***
	0.165	*0.467*	*0.355*	*0.195*	*0.221*	*0.240*	*0.262*	*0.275*
N	46,827	46,827	46,827	46,827	46,827	46,827	46,827	46,827
$\hat{\alpha}$			40.06***					
Log-Likelihood	-4,989.07	-15,242.52	-6,839.61	-7,879.36	-8,967.57	-9,554.84	-10,103.04	-10,511.68

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level

Table 24. Quantiles for Counts

Quantiles	**0.025**	**0.500**	**0.700**	**0.800**	**0.900**	**0.950**	**0.990**	**0.999**	**0.9999**
IMMIGRANT	-0.038	-0.319	-0.422**	-0.440***	-0.498***	-0.599***	-0.393**	-0.609*	-0.321
	0.061	*0.981*	*0.199*	*0.152*	*0.141*	*0.188*	*0.206*	*0.326*	*0.251*
MALE	0.054*	0.235***	0.423***	0.449***	0.532***	0.548***	0.285**	-0.206	-0.514***
	0.030	*0.052*	*0.063*	*0.063*	*0.081*	*0.106*	*0.121*	*0.209*	*0.157*
AGE 26 – 35	-0.083***	-0.371***	-0.580***	-0.594***	-0.807***	-2.158***	-0.782***	-0.733*	0.121
	0.035	*0.065*	*0.079*	*0.079*	*0.130*	*0.242*	*0.225*	*0.419*	*0.155*
AGE 36 – 45	-0.129***	-0.721***	-1.080***	-1.134***	-1.363***	-3.067***	-1.112***	-0.604	0.075
	0.044	*0.193*	*0.098*	*0.090*	*0.136*	*0.127*	*0.216*	*0.465*	*0.255*
AGE 46 – 55	-0.187	-0.947***	-1.537***	-1.600***	-1.827***	-3.517***	-1.703***	-1.916***	-1.192***
	0.164	*0.138*	*0.152*	*0.180*	*0.155*	*0.133*	*0.242*	*0.360*	*0.227*
AGE 56 – PLUS	-0.229***	-1.519	-2.714	-2.790***	-3.129***	-4.749***	-6.346***	-2.775***	-2.272***
	0.069	*4.427*	*5.025*	*0.294*	*0.458*	*0.152*	*0.223*	*0.285*	*0.142*
DEPRIVATION	0.002	0.029***	0.055***	0.053***	0.062***	0.088***	0.080***	0.081***	0.073***
	0.005	*0.011*	*0.013*	*0.013*	*0.015*	*0.020*	*0.026*	*0.044*	*0.025*
URBAN	0.001	0.023	0.128	0.113	0.163*	0.182	-0.007	-0.208	-1.165***
	0.032	*0.073*	*0.117*	*0.086*	*0.096*	*0.116*	*0.149*	*0.203*	*0.149*
INNER CITY	-0.014	0.024	-0.058	-0.004	0.040	0.024	0.150	0.152	0.269
	0.063	*0.097*	*0.119*	*0.115*	*0.141*	*0.180*	*0.261*	*0.344*	*0.209*
NORTH EAST	0.044	0.387*	0.466	0.493**	0.576***	0.642**	0.071	-0.299	-0.127
	0.074	*0.216*	*0.335*	*0.197*	*0.209*	*0.264*	*0.259*	*0.357*	*0.227*
NORTH WEST	-0.025	0.224	0.200	0.259	0.308*	0.382	0.237	-0.129	-0.059
	0.080	*0.212*	*0.329*	*0.185*	*0.183*	*0.249*	*0.259*	*0.262*	*0.223*
YORKSHIRE	0.012	0.269	0.334	0.374*	0.426**	0.568**	0.327	-0.008	0.037
	0.073	*0.227*	*0.362*	*0.193*	*0.194*	*0.266*	*0.267*	*0.338*	*0.211*
EAST MIDLANDS	-0.018	0.273	0.478	0.515***	0.620***	0.737***	0.612*	0.260	0.799***
	0.097	*0.229*	*0.332*	*0.184*	*0.188*	*0.251*	*0.329*	*0.309*	*0.211*
WEST MIDLANDS	0.025	0.230	0.346	0.379**	0.488***	0.627**	0.578*	0.601	1.824***
	0.077	*0.223*	*0.334*	*0.190*	*0.188*	*0.251*	*0.294*	*0.393*	*0.194*
EAST OF ENGLAND	-0.011	0.048	0.177	0.229	0.258	0.373	0.167	0.253	1.582***
	0.073	*1.104*	*0.335*	*0.189*	*0.188*	*0.245*	*0.271*	*0.335*	*0.250*
SOUTH EAST	0.013	0.256	0.384	0.412**	0.469**	0.659**	0.397	0.610	2.225***
	0.107	*0.341*	*0.328*	*0.189*	*0.189*	*0.260*	*0.290*	*0.516*	*0.328*
SOUTH WEST	-0.107	0.163	0.242	0.284	0.368*	0.499*	0.273	0.209	2.475***
	7.310	*0.402*	*0.381*	*0.205*	*0.194*	*0.263*	*0.272*	*0.459*	*0.271*
WALES	0.049	0.340	0.393	0.396**	0.480**	0.725***	0.295	1.236***	2.153***
	0.066	*0.224*	*0.354*	*0.199*	*0.203*	*0.275*	*0.275*	*0.615*	*0.281*
CONSTANT	-0.512***	-2.313***	-3.479***	-3.449***	-3.430***	-1.974***	0.225	2.624***	3.463***
	0.074	*0.216*	*0.338*	*0.195*	*0.195*	*0.327*	*0.329*	*0.410*	*0.302*
N	46,827	46,827	46,827	46,827	46,827	46,827	46,827	46,827	46,827

Standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level

Table 25. Hurdle Poisson for Censored Counts

	Hurdle	Zero Trunc Poisson	Zero Truncated Censored Poisson	
			5	10
IMMIGRANT	-0.360***	0.047	0.002	-0.026
	0.110	*0.512*	*0.189*	*0.205*
MALE	0.430***	-1.106***	-0.400***	-0.446***
	0.058	*0.243*	*0.098*	*0.114*
AGE 26 – 35	-0.616***	0.018	-0.175	-0.128
	0.076	*0.290*	*0.130*	*0.145*
AGE 36 – 45	-1.153***	0.498	-0.022	0.005
	0.083	*0.348*	*0.130*	*0.147*
AGE 45 – 56	-1.566***	-0.275	-0.491***	-0.547***
	0.101	*0.551*	*0.179*	*0.219*
AGE 56 – PLUS	-2.765***	-0.778*	-0.693***	-0.652***
	0.111	*0.410*	*0.221*	*0.282*
DEPRIVATION	0.052***	-0.032	0.033	0.033
	0.012	*0.048*	*0.021*	*0.024*
URBAN	0.123	-0.362	-0.268**	-0.269**
	0.078	*0.326*	*0.126*	*0.145*
INNER CITY	0.032	-0.077	0.029	0.062
	0.103	*0.274*	*0.179*	*0.196*
NORTH EAST	0.530***	-0.035	-0.504	-0.345
	0.159	*0.441*	*0.323*	*0.365*
NORTH WEST	0.259*	-0.023	0.024	0.026
	0.149	*0.317*	*0.263*	*0.276*
YORKSHIRE	0.388**	0.513	0.238	0.327
	0.153	*0.352*	*0.254*	*0.275*
EAST MIDLANDS	0.523***	0.673	0.080	0.127
	0.147	*0.419*	*0.258*	*0.275*
WEST MIDLANDS	0.437***	0.954***	0.303	0.480*
	0.149	*0.369*	*0.250*	*0.265*
EAST OF ENGLAND	0.224	0.985*	0.012	0.207
	0.152	*0.546*	*0.278*	*0.301*
SOUTH EAST	0.388**	0.556	0.153	0.157
	0.156	*0.512*	*0.259*	*0.278*
SOUTH WEST	0.333**	0.577	0.007	0.210
	0.157	*0.450*	*0.283*	*0.308*
WALES	0.446***	1.426***	0.089	0.267
	0.158	*0.478*	*0.273*	*0.301*
CONSTANT	-3.365***	1.018**	0.360	0.473
	0.161	*0.513*	*0.288*	*0.315*
N	46,827	1,190	1,190	1,190
Log-Likelihood	-4,988.90	-4,361.55	-1,326.98	-1,751.82

Quasi-Maximum Likelihood standard errors are presented in *italics*.
(***) denotes statistical significance at 1% significance level,
(**) denotes statistical significance at 5% significance level,
(*) denotes statistical significance at 10% significance level

Appendix: A Hurdle-Poisson Model for Censored Counts

This model combines results from hurdle models and censored models for counts as presented by Mullahy (1986) and Terza (1985), respectively. The hurdle part of the model recognizes that the binary outcome (zeroes or positives) is generated by a probability distribution appropriate for binary models, while the counts are generated by a truncated at zero distribution appropriate for count data. However, this model is modified to take into account that once the hurdle is crossed the probability function that has support only over the positive counts is censored at C. According to this, the probability of a zero, the probability of a positive but uncensored integer, and the probability of a censored outcome are given by,

$$
\begin{aligned}
&\Pr(y=0) = f_1(0),\\
&\Pr(y=k|0<y<C) = (1-f_1(0))\left(\frac{f_2(y)}{1-f_2(0)}\right),\\
&\Pr(y \geq C) = 1 - f_1(0) - (1-f_1(0))\left(\frac{f_2(1)-f_2(2)-f_2(3)\ldots f_2(C-1)}{1-f_2(0)}\right),
\end{aligned}
$$

where $1-f_2(0)$ is used as a normalization to account for the zero truncation. In the present study we assume that both $f_1(.)$ and $f_2(.)$ are Poisson distributed. In a regression framework, conditional on a set of characteristics $\mathbf{x}_i$, with $i = 1, \ldots, n$, which is assumed to be common in both processes, $f_1(.)$ and $f_2(.)$ follow the Poisson distribution with $\lambda_1 = \exp(\mathbf{x}_i'\boldsymbol{\beta})$ and $\lambda_2 = \exp(\mathbf{x}_i'\boldsymbol{\gamma})$. The likelihood function is given by,

$$
\begin{aligned}
L(\beta,\gamma) = &\prod_{i=1}^{n} f_1(0)^{(y=0)} \times \left[\left(\frac{1-f_1(0)}{1-f_2(0)}\right) f_2(y)\right]^{(0<y<C)}\\
&\times \left[1 - f_1(0) - \left(\frac{1-f_1(0)}{1-f_2(0)}\right)(f_2(1)+f_2(2)+f_3(3)+\ldots f_2(C-1))\right]^{y\geq C}\\
= &\prod_{i=1}^{n} (e^{-\lambda_1})^{(y=0)} \times \left[\left(\frac{1-e^{-\lambda_1}}{1-e^{-\lambda_2}}\right)\frac{e^{-\lambda_2}\lambda_2^{y_i}}{y_i!}\right]^{(0<y<C)}\\
&\times \left[1-e^{-\lambda_1} - \left(\frac{1-e^{-\lambda_1}}{1-e^{-\lambda_2}}\right)\left[e^{-\lambda_2}\left(\lambda_2 + \frac{\lambda_2^2}{2} + \frac{\lambda_2^3}{3!} + \ldots + \frac{\lambda_2^{C-1}}{(C-1)!}\right)\right]\right]^{y\geq C},
\end{aligned}
$$

which collapses to the standard Censored Poisson model if $\lambda_1 = \lambda_2$. Now, once we multiply and divide the second term by e^{λ_1}, the log likelihood is the following:

$$\ln L = \sum_{i=1}^{n} (y=0)(-\lambda_1) + (0<y<C)\left[\ln(1-e^{-\lambda_1}) - \ln(e^{\lambda_2}-1) - \ln(y_i!) + y_i \ln\lambda_2\right]$$
$$+ (y \geq C)\ln\left[(1-e^{-\lambda_1}) - \left(\frac{1-e^{-\lambda_1}}{e^{\lambda_2}-1}\right)\left(\lambda_2 + \frac{\lambda_2^2}{2} + \frac{\lambda_2^3}{3!} + \ldots + \frac{\lambda_2^{C-1}}{(C-1)!}\right)\right],$$

which can be further simplified as,

$$\ln L = \sum_{i=1}^{n} (y=0)(-\lambda_1) + (y>0)\ln(1-e^{-\lambda_1}) + (0<y<C)\left[-\ln(e^{\lambda_2}-1) - \ln(y_i!) + y_i \ln\lambda_2\right]$$
$$+ (y \geq C)\ln\left[1 - \left(\frac{1}{e^{\lambda_2}-1}\right)\left(\lambda_2 + \frac{\lambda_2^2}{2} + \frac{\lambda_2^3}{3!} + \ldots + \frac{\lambda_2^{C-1}}{(C-1)!}\right)\right].$$

From the last expression it is clear that the log likelihood function is separable. This simplifies the estimation procedure as we can separately maximize the likelihood part of the binary outcome, using all observations, and the likelihood part of the zero truncated censored counts using only the positive counts. Turning the last term into a fraction with common denominator, and separating it into two logs we can finally rewrite the likelihood function as,

$$\ln L = \sum_{i=1}^{n} (y=0)(-\lambda_1) + (y>0)\ln(1-e^{-\lambda_1})$$
$$- (y>0)\ln(e^{\lambda_2}-1) + (0<y<C)\left[-\ln(y_i!) + y_i\ln\lambda_2\right]$$
$$+ (y \geq C)\ln\left[e^{\lambda_2} - 1 - \lambda_2 - \frac{\lambda_2^2}{2} - \frac{\lambda_2^3}{3!} - \ldots - \frac{\lambda_2^{C-1}}{(C-1)!}\right].$$

Maximum likelihood estimation follows using numerical algorithms, such as the Newton-Raphson.