Criminal Behavior and Immigration

An Application of some Estimators for Under-reported Outcomes using the Offending, Crime, and Justice Survey

Georgios Papadopoulos*

Department of Economics, University of Essex

January 2011

Abstract

This paper mainly studies the individual level relationship between immigration and property crime in England and Wales. For this purpose, the Offending, Crime, and Justice Survey is used, a representative national survey of self-reported crime. Models that account for under-reporting are used, as this is the major concern in self-reports. The results of these models indicate that under-reporting of crime in the OCJS is considerably large. Moreover, they also indicate that under-reporting is not constant but it rather depends on respondents' characteristics. However, our findings suggest that, if anything, immigrants tend to under-report by less than natives. Binary choice models reveal that, after controlling for under-reporting and for basic demographic characteristics, the probability of committing a crime is lower for immigrants, but the difference is statistically insignificant. This finding is evident in count data models as well, as being an immigrant (insignificantly) decreases the mean number of crimes. Furthermore, violent crime results are in line with the findings of property crime as the immigrant-crime association is also negative. Interestingly, a decomposition of immigrants by regions reveals that different regions attract immigrants of different criminal behavior, or that immigrants adapt differently across regions. Finally, the results also show that immigrants of different ethnic groups exhibit disparate criminal behavior.

Keywords: Crime; Immigration; Self-reports; Under-reporting; MisProbit; NB2-Logit.
JEL Classification Numbers: K42, J15, J22, C25, C51

*Address: Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. E-mail: *gpapad@essex.ac.uk*.
I am grateful to João Santos Silva for his excellent guidance and support. I would also like to thank, Holger Breinlich, Zelda Brutti, Ken Burdett, Tim Hatton, Mariña Fernádez Salgado, Michail Veliziotis, Gianluigi Vernasca, participants in the Internal and the Research Strategy Seminar of the Department of Economics, University of Essex, and participants in the RES Sixth PhD Meeting, SAEe 2010, EEA 2010, WPEG 2010, the "Economics of Culture, Institutions, and Crime" conference at FEEM, Milan, and the Crime Surveys User Meeting, ESDS, London, for their helpful comments. The empirical analysis is implemented in Stata/SE and TSP econometric software. Finally, the data used in this project are sponsored by the Home Office and provided by the UK Data Archive.

1 Introduction

Debates on the relationship between immigration and crime date back at the very beginning of the twentieth century (see, for example, Hourwich, 1912), when there was a big influx of European and Canadian citizens into the United States. It seems that the native born population of countries that sustained heavy migration inflows always developed hostile feelings against the foreigners. This can be attributed to the fact that natives always feared that immigrants could take away their jobs and deteriorate several problems of the host countries, including crime. Although the bulk of the media in the host countries supported and even strengthened this negative perception against immigrants, researchers' community often concluded the opposite. Many found evidence that immigrants seem to commit even less crimes than natives, even though they usually encounter unfavorable circumstances, such as blocked opportunities, or acculturation problems (Tonry, 1997, Hagan and Palloni, 1998, and Mears, 2001).

The present study attempts to investigate the relationship between immigration and crime in England and Wales.[1] In spite of the fact that the immigration-crime link is such a controversial subject, the immigration-crime association is generally overlooked by the research community relative to other aspects of crime. To my knowledge, apart from the very recent study by Bell, Machin and Fasani (2010), there is no study that investigates whether such a relationship exists in the UK.[2]

As figure 1.1 demonstrates, England and Wales have recently experienced a steady increase in immigration stock relative to the total population.[3] It is notable, as Hatton and Tani (2005) show, that immigration net flows are responsible for about one half of the population growth during the previous decade. Moreover, Hatton (2005) explains that after reaching the minimum at 1993-1995, immigrant net flows increased until 2000, reaching the figure of 100,000 individuals per year. Recently, immigration rates increased even more,

[1]Scotland and Northern Ireland are excluded from the survey used to investigate this research question because of their separate criminal and justice system, which generates not comparable crime statistics.

[2]However, there is a crucial difference between the present study and the study of Bell, Machin and Fasani (2010); that is, the present study looks at this relationship from a micro perspective, examining whether immigrants are more prone to criminal activities than natives, whereas their study focuses on the effect of two waves of large inflows of immigrants on crime rates.

[3]Although the OECD provides statistics for immigration rates until 2005, estimates from the Labour Force Survey show that immigrant population have kept increasing during the last 3 years.

because of the heavy influx of around 560,000 Eastern European workers between 2004 and 2006.[4]

Following this increase in foreign population, immigration in the UK became a very controversial subject, and one of the "hottest" topics in political agenda. During the last two decades natives have developed negative perceptions against their immigrant counterparts, with regard to labor market outcomes, cultural issues, and crime. This hostile tendency is quite clear if we look at the UK sample of two very important attitudes surveys, the European Social Survey (ESS) of 2002, and the International Social Surveys (ISS) of 1995 and 2003, where questions related to immigration and crime are included.

According to the ESS (see Table 1.1) there is a clear tendency towards the perception that immigrants have worsened UK's crime rates. Even more interesting findings come from the ISSs (Table 1.2). Although in ISS of 1995 only 26% of the respondents believe that immigrants increase crime rates (agree and strongly agree), this figure increases to around 40% in 2003. These findings coincide with the increase in immigrant population. It is quite interesting that natives developed negative perceptions against immigrants despite the fact that crime rates (at least for total crime) started falling after 1995 (see, figure 1.1 and 1.2. Also see, Smith, 2006, and Kershaw and al, 2008, p.2). To reinforce these findings, an Ordered Probit regression model is presented in Table 1.3, where a simple dummy for year 2003 aims to capture the evolution in natives' attitudes, once we have pooled the data from ISS 1995 and ISS 2003. This is done because we recognize that the sample in 1995 might differ in many aspects from the sample in 2003. It is clear that even after controlling for some basic characteristics correlated with respondents' attitudes, such as education, party affiliation, income, gender, and age, moving from 1995 to 2003 strongly increased the sentiment that immigrants increase crime rates.[5]

But which are the theoretical reasons that link immigration with crime? There are two

[4] This evidence comes from the Worker Registration Scheme and National Insurance Number applications (see, Gilpin at al, 2006, Blanchflower at al, 2007, and Lemos and Portes, 2008).

[5] The base group for the variable "party" is "left-wing". For "education" the base group is "lower education". The marginal effects on the "year 2003" dummy, which are not presented here but are available from the author upon request, show that the probability of responding with "agree that immigrants increased crime rates" increased by 7.3 percentage points from 1995 to 2003, and the probability of responding with "strongly agree…" by 5.9 percentage points. Both of them are statistically significant at 1% significance level.

distinct effects of immigration on crime. The first one, that I call the "aggregate" effect or macro effect, states that immigration inflows are related to economy's crime rates as they: 1) can affect aggregate outcomes of the domestic economy, such as wages and unemployment, and 2) may impose cultural conflicts and social disorganization according to criminological theories (see for example, Martinez and Lee, 2000). The second one, which I call the "individual" or micro effect, is the direct effect. This states that immigrants are more or less crime-prone than natives because of some observed or unobserved characteristics specific to them. This paper attempts to shed light on the direct relationship by investigating the question: *Do immigrants commit more or less crime than natives and why?* It should be also stressed that this work focuses on property crime, which can be better explained by economic theory, as opposed to violent crime, which is better explained by psychological factors rather than material needs. Therefore, psychological theories developed basically by criminologists, rather than economic theory, would stand better to explain violent crime.[6] In the next section a simple economic model of crime is presented to investigate the direct immigration-crime relationship.

For the purposes of the empirical analysis individual level crime data are needed. A first choice would be to use recorded by police crime and compare the crime records of immigrants to those of natives. However, data on recorded crime in England and Wales are very poor in terms of information provided, making them inappropriate for an individual empirical investigation. Particularly relevant to this research is the fact that these data provide no information on the immigration status of criminals. In Section 4, other shortcomings of this kind of data are explained. In another direction, self-reports on crime can be used, a practice that is very common among criminologists and sociologists (see, Junger-Tas and Marchall, 1999). Therefore, in the present study the Offending, Crime and Justice Survey (OCJS) of 2003 is used, a nationally representative survey that asks people in England and Wales about their experiences and attitudes towards criminal activities (Hamlyn at al, 2003). Not only does this kind of surveys reveal in some extent unreported/unrecorded crime, but it also provides a rich set of respondents' attributes that allows identification of

[6]However, as the explanatory variables used to the empirical analysis may also determine the decision to commit violent crimes, the results of a violent crime model are presented in subsection 7.3.

those characteristics that lead to criminal behavior.

Of course, to identify these relationships it is required that respondents truthfully reveal their criminal activity. Nevertheless, reliability of self-reports on crime is a major concern as many individuals may be reluctant to provide such sensitive information. Therefore, under-reporting is a major concern, although nowadays many techniques are used to improve the reliability and validity of these data (Thornberry and Krohn, 2000). From the econometric point of view, estimators that ignore under-reporting are inconsistent (see for example, Hausman, Abrevaya, and Scott-Morton, 1998, for binary outcome models, and Winkelmann and Zimmermann, 1993, for count data models). Therefore, it is highly possible that they provide misleading estimates for the coefficients of interests. This problem becomes even more salient if differences in respondents' characteristics are associated with different reporting behavior, and more importantly, if immigrants reporting behavior differs from natives' one. However, appropriate econometric models incorporating this problem can be developed and applied. These models consistently estimate the determinants of true crime, using only self-reported figures of crime. This is the subject of Section 5.

Initially, we treat immigrants as a homogeneous group of people. However, this may not be proper for plenty of reasons. For example, immigrants of different ethnic backgrounds might be very different from each other. Furthermore, location of immigrants is not randomly assigned, but it is a rather complicated process that depends on many factors. For example, if immigrants try to match their abilities with the opportunities that each area provides, more crime-prone immigrants would decide to reside in areas that offer more criminal opportunities. Or, as the location of immigrants also depends on central decisions, it might be that different kind of immigrants are located in areas characterized by different socioeconomic features. For example, following the 1999 Immigration and Asylum Act, asylum seekers were located by the National Asylum Support Service in specific areas, London being excluded (see, Bell, Machin, Fasani, 2010). Given the facts above, the estimated effect of being an immigrant on crime might be misleading and therefore, the effect of immigration on crime is also investigated once we decompose immigrants by ethic group and location.

The remainder of this paper is organized as follows. Section 2 puts the individual decision to commit property crimes in a simple economic framework of individual supply of crime. Utilizing this simple economic model, it also investigates the individual relationship between immigration and property crime. In Section 3 a very brief review of studies on this topic is presented. In Section 4 some methodological issues of self-reports are discussed. Section 5 offers a presentation of the econometric models that are more appropriate in the presence of under-reporting. Section 6 discusses the data and the variables and offers some basic descriptive statistics. The main results follow in Section 7. In the same section robustness of these results is checked. Moreover, this section also investigates whether immigrants of different ethic status or immigrants risiding in different locations are associated with differential criminal behavior. Finally, discussion of the empirical results follows in Section 8, and Section 9 concludes.

2 An Economic Model of Property Crime

As discussed in the introduction, this study investigates whether immigrants are different from natives with regard to their behavior towards criminal activities, and particularly towards property crime. Therefore, in this section the individual relationship between immigration and crime is examined. We start with a general examination of a simple model of property crime. In the next subsection we examine how immigration status is associated with this model, and consequently what this model predicts about criminal activity of immigrants compared to natives.

Generally, the economic theory of crime is based on the idea of the rational individual who chooses how to allocate his time between legitimate and illegitimate activities so as to maximize his personal expected utility.[7] As there is a probability of apprehension, the final outcome of the criminal act is uncertain.

Becker (1968) has offered the first prominent paper to incorporate economic theory on criminal behavior. However, in this early work illegal and legal activities were considered as mutually exclusive. Few years later, Ehrlich (1973), in a cornerstone work, relaxed this assumption so as individuals are utility maximizers who allocate their time between crime

[7] I describe the individual as a “he” just because crime is generally considered as an activity associated with males.

and work. In another vital work, Block and Heineke (1975) criticize the two previous works on the grounds that they treat crime and punishment outcomes as if they can be always represented by their pecuniary equivalents. Using a more general, multiattributed utility function, they show that the results of Becker's and Ehrlich's works hold only under very special conditions, and that determining the supply of property crime is a harder task that needs further assumptions.[8] From then on, many other economic theoretical models have been developed[9] and tested using micro or macro data.[10] It needs to be stressed that in general, property crime fits better in the economic models of crime, since violent crimes can be considered as non market activities that are primarily motivated by hate or passion (Ehrlich, 1973).[11]

Economic models of crime have been extended beyond the classical theory of crime deterrence, particularizing in examining relationships such as investment in human capital and crime (Lochner, 2004, Lochner and Moretti, 2004), inequality and crime (Chiu, Madden, 1998), the effect of economic incentives on crime (Machin and Meghir, 2004), crime and unemployment (Burdett, Lagos and Wright, 2003), crime and social interactions (Glaeser, Sacerdote and Scheinkman, 1996), etc. Nevertheless, there is no theoretical framework that investigates the relationship between immigration and crime. For this reason, the following subsection presents a simple model that incorporates immigration with the purpose of demonstrating why immigrants might exhibit different criminal behavior than natives.

To make it as simple as possible, the following model is a one period model under uncertainty that borrows features from Ehrlich (1973), and Lochner and Moretti, (2001). This model is by no means a complete investigation of criminal behavior, but it illustrates quite well why someone would expect differences in participation rates of illegitimate activities between immigrants and natives.

[8] The utility function is given as consisting of three attributes, $U(L, T, W)$, (where L and T are, time spent on legal and illegal activities respectively, and W represents wealth), rather than a wealth only function.

[9] Cameron (1988) and Eide (1999) are good surveys on this topic. Freeman (1999) is also an excellent survey that discusses many aspects surrounding the economic theory of crime from both a theoretical and an empirical perspective.

[10] See, for example, Sjoquist (1973), Woplin (1978), Witte (1980), Myers (1983), Reilly and Witt (1996), Cornwell and Trumbull (1994), and Kelly (2000), to mention only a few.

[11] However, as will be clear later, some factors determining property crimes, such as probability of apprehension, severity of punishment and risk aversion, are directly associated with violent crime as well.

Consider a rational individual who, after receiving the initial endowment z, optimally decides how to allocate his total time available, τ, between criminal activity, τ^i, and work, τ^l. We assume that leisure time, where the individual consumes all his outcomes, is constant and therefore does not affect the results of the model.[12] Although in general z can represent other individual characteristics such as, age, gender, parental features, respondent's location features, etc., in this model z is an indicator variable that determines immigration status. In turn, z is assumed to affect most of the parameters of this model.

Uncertainty is incorporated in the model because of two reasons. First, there is a probability of apprehension, $\pi(\tau^i, z)$, in case the individual is involved in criminal activities, which is an increasing function of τ^i. Second, legal outcomes are also not certain because there is a probability of unemployment, $\mu(z)$, which is assumed to be given exogenously at the beginning of the period.

In the legal sector the individual faces the legal wage $w(\tau^l, z)$. This legal wage depends positively on τ^l, such that $w'(\tau^l, z) > 0$, and $w''(\tau^l, z) < 0$.[13] Apart from legal opportunities, the individual also faces illegal opportunities, given by $k(\tau^i, z)$, which consists of financial and psychic outcomes measured in their pecuniary equivalent.[14] Apart from financial and psychic gains, $k(\tau^i, z)$ also includes some costs (measured in their pecuniary value) associated with a crime, such as bad reputation, compunction, regrets, psychic uneasiness, etc. The costs of trial, conviction and punishment, are not included in these costs but, as we will see shortly, they will be introduced as distinct components of the utility function. This criminal "wage" depends continuously on τ^i, such as $k'(\tau^i, z) > 0$, and $k''(\tau^i, z) < 0$.[15] For simplicity, in case the individual is actually unemployed, the only source of income is the income he receives from illegal activities. Thus, when the individual enters the market he considers a continuous set of illegal opportunities. We assume that illegal opportunities that pay high pecuniary returns require a lot of time in illegal sector

[12]Without loss of generality, τ can be considered as the time available for allocation between the legal and illegal activities after extracting leisure time from total available time.

[13]$w'(\tau^l, z)$ is the rate of change of the legal wage, $dw(\tau^l, z)/\tau^l$, and $w''(\tau^l, z)$ is the second derivative $d^2w(\tau^l, z)/(d\tau^l)^2$. The latter can be assumed to be negative as productivity and efficiency may decrease as more time is spent on work.

[14]By pecuniary equivalent we mean the amount of money that someone is willing to pay in order to get this gain or to avoid a cost.

[15]Where again, $k'(\tau^i, z) = dk(\tau^i, z)/d\tau^i$, and $k''(\tau^i, z) = d^2k(\tau^i, z)/(d\tau^i)^2$.

and that they also involve higher psychic costs.[16]

There are two states of nature, the good State A, where someone is employed with probability $1 - \mu(z)$, and the bad State B, where someone is unemployed with probability $\mu(z)$. The corresponding expected returns from legal and illegal actions are the following,

$$\begin{aligned} y_a &= w(\tau^l, z) + k(\tau^i, z), \\ y_b &= k(\tau^i, z), \end{aligned} \tag{2.1}$$

where y_a is associated with State A, and y_b with State B. Therefore, the expected utility once consuming y_a, y_b, (without considering potential punishment) is given by,

$$(1 - \mu(z))\, u(y_a) + \mu(z) u(y_b) = (1 - \mu(z))\, u\left(w(\tau^l, z) + k(\tau^i, z)\right) + \mu(z)\, u\left(k(\tau^i, z)\right) \tag{2.2}$$

with $u'(y_i) > 0$, and $u''(y_i) < 0$, where $i = (a, b)$.

Moreover, crime is a risky action. Thus, if someone is involved in criminal activities he faces a probability of arrest, $\pi(\tau^i, z)$, as described before. We assume that this probability increases with time spent on illegal sector, so that $\pi'(\tau^i, z) > 0$, but the second derivative can be either positive or negative.[17] Despite the fact that arrested people are not always convicted, here we assume that if arrested, conviction, and thus, punishement is certain. Nothing is lost from this simplification since it can be shown that it does not affect the implications of the model. Conviction occurs at the end of the period, where the individual receives a punishment $P(\tau^i, z)$, pecuniary or not pecuniary such as imprisonment, with $P'(\tau^i, z) > 0$ and $P''(\tau^i, z) > 0$.[18] According to the above, the present value of expected future punishment is given by,

$$\Pi(\tau^i, z) = \rho(z)\pi(\tau^i, z)P(\tau^i, z) \tag{2.3}$$

[16] Therefore, regardless of psychic costs, someone who pursuits high returns to illegal actions can either commit many crimes, or one high value crime which requires much time spent in the illegal sector though. This can be the case, as this type of property crimes requires much time for organization, preparation, etc.

[17] It could be negative, since self-protection improves as people spend more time in criminal activities. On the other hand, it could be positive as well, as more time in the illegal sector allows the law enforcement to acquire more evidence against the criminal, which increases the probability of apprehension.

[18] Any kind of punishment, as in the cases on Becker (1968) and Ehrlich (1973), is measured in its monetary equivalent. Here it is assumed that the more serious the crime the stricter the punishment becomes.

where $\rho(z)$ discounts punishment since it occurs at the end of the period. For simplicity, expected punishment is measured in utility terms as in Lochner and Moretti, (2001).[19] Henceforth, z is omitted from the equations for brevity.

Given all the above, the total expected utility received by both legal and illegal activity from both states is the following,

$$U(\tau^i, \tau^l) = (1-\mu)\, u\,(y_a) + \mu\, u\,(y_b) - \rho\, \pi(\tau^i) P(\tau^i). \tag{2.4}$$

Thus, the problem of the individual is to allocate his available time between legal and illegal activities in order to maximize (2.4) subject to the time constraints,

$$\tau = \tau^i + \tau^l, \qquad \text{and,} \qquad \tau^i \geq 0, \tau^i \geq 0 \tag{2.5}$$

After substituting $\tau^l = \tau - \tau^i$, and (2.1) into (2.4), (which turns it into a one variable maximization problem), the first order condition for an interior solution of committing at least one crime is given by,

$$\frac{dU\,(\tau^i)}{d\tau^i} = 0,$$

$$\Rightarrow (1-\mu) u'(y_a) \left[k'(\tau^i) - w'(\tau^l)\right] + \mu\, u'(y_b) k'(\tau^i) = \rho \left[\pi'(\tau^i) P(\tau^i) + \pi(\tau^i) P'(\tau^i)\right], \tag{2.6}$$

where $w'(\tau^l)$ can be considered as the opportunity cost of crime of not spending the extra time $d\tau^i$ on the legal sector. The sufficient condition for a global maximum is given by,

$$\Delta = (1-\mu) \left[u''(y_a)(k'-w')^2 + u'(y_a)(k''+w'')\right] \\ + \mu \left[u''(y_b)(k')^2 + u'(y_b)k''\right] - \rho(\pi'' P + 2\pi' P' + \pi P'') < 0. \tag{2.7}$$

[19] That this future potential punishment is measured in utility terms has as implication that punishment is separable from (2.2). Otherwise, this future punishment should have been incorporated in the indirect utility functions in the same manner as in (2.2). In that case, there should have been four mutually exclusive states, for employed and not arrested, unemployed and not arrested, employed and arrested, and unemployed and arrested, as described in Ehrlich (1973). This would result in four mutually exclusive utility outcomes, each associated with the probability of the state of nature to be observed and the total expected utility would have been, $U(\tau^i) = \left(1-\pi(\tau^i)\right)(1-\mu) u\left(w(\tau^l) + k(\tau^i)\right) + \left(1-\pi(\tau^i)\right) \mu \cdot u\left(k(\tau^i)\right) + \pi(\tau^i)(1-\mu) u\left(w(\tau^l) + k(\tau^i) - \rho P(\tau^i)\right) + \pi(\tau^i)\mu \cdot u\left(k(\tau^i) - \rho P(\tau^i)\right)$ rather than the simpler function (2.4).

According to (2.6), and given that (2.7) holds, the individual devotes some time to criminal activities, if the expected marginal net gains from spending the extra time $d\tau^i$ on the illegal sector are equal to the potential punishment of this extra criminal time. Since the term on the right hand side is weakly positive, it is required that the marginal criminal returns are at least higher than the marginal legal returns ($k'(\tau^i) > w'(\tau^l)$), or at least that $(1-\mu)u'(y_a)[k'(\tau^i) - w'(\tau^l)] < \mu u'(y_b)k'(\tau^i)$, in case $k'(\tau^i) \leq w'(\tau^l)$, so that the expected net utility gains from both states of nature are positive. This is because crime is a risky action that involves losses in the case of a potential future punishment. Thus, the term on the right hand side can be considered as the extra marginal compensation required for crimes to be committed.

Since the criminal wage rate is in general quite small relatively to the legal wage rate for most property crimes, and if we consider that for most people the criminal wage further decreases by the psychic costs associated with a crime, the corner solution where someone allocates all his time in legal actions is highly possible.[20] Moreover, property crimes that pay a quite high marginal return are also very rare, as according to our assumptions, crimes that pay high returns require plenty of time which in turn increases the risk of apprehension and the severity of punishment.[21] Also, crimes that pay a lot involve much higher psychic costs than psychic gains for most people, so that $k(\tau^i)$ is not large enough. Finally, we must also consider that for many individuals $k(\tau^i)$ is quite small just because these individuals do not exhibit strong criminal ability. All the above are possible reasons to explain why crime is such a rare event.[22]

A main point of this equation is that, starting from an equilibrium where the individual spends some time on the illegal sector, the better the opportunities in legal sector, expressed as higher $w'(\tau^l)$, the higher the opportunity cost of crime is. Therefore, holding everything else constant, the participation in illegal activities will decrease. The opposite holds for the marginal return to crime. As it becomes higher relative to the marginal return to legal activities, the individual is better off if he allocates more time to criminal activities than

[20]For most people, property crimes would include more psychic costs because of regret, bad reputation, etc, rather than psychic gains because of possible satisfaction.

[21]Think for example bank robberies, or car thefts. Although the crimes itself may not need so much time, we can assume that they need a lot of preparation, which is also measured in τ^i.

[22]In my sample the proportion of people that have reported a crime during the last period is just 5%.

before.

Nevertheless, starting from an interior solution for crime, the effect of an increase in unemployment rate is less obvious. This happens because the increase in unemployment probability lowers the utility obtained from State A but increases the utility obtained from State B. The comparative static analysis shows that $d(\tau^i)^*/d\mu = \left[u'(y_a)(k'-w') - u'(y_b)k'\right]/\Delta$, which is positive only if $k'/(k'-w') > u'(y_a)/u'(y_b)$, (where both ratios are higher than one) so that the relative marginal return from spending some extra time on crime is higher than the relative marginal utility obtained from both states (so, it depends on the weight the individual puts on utilities relative to the marginal return).

The effect of all the components of the potential punishment is also the expected one. For example, following an exogenous increase in the probability of punishment, the marginal return to criminal activities must also go up to compensate this increase in potential punishment. Otherwise the individual would decrease τ^i. The same will be the effect if there is an exogenous increase in the severity of punishment. In addition, risk attitudes, which can be expressed through the discount factor or the curvature of the utility curves, are quite important. As someone becomes riskier or more impatient, he discounts future potential punishment more heavily (lower $\rho(z)$). In this way, the marginal return from crime must be higher for a less risky or a more patient individual, since he puts much weight on the consequences of a possible future apprehension. Moreover, as y goes up, $u'(y)$ decreases more for a more risk averse individual, accordingly decreasing the left hand side of (2.6), resulting in a higher extra compensation required for the less risky person to participate on the illegal sector.

2.1 Immigration and Crime

What could this simple model tell us about immigrants' behavior towards crime? Since z determines whether someone is an immigrant or a native, immigration status affects the first order condition (2.6) through many channels. Immigrant population is not a homogeneous group of people, as it includes individuals with very different ethnic backgrounds. However, they exhibit some common features that distinguish them from natives. There are some reasons to believe that there may be a positive relationship between immigration status and crime, and some others that would suggest a negative link.

We would expect a positive link since immigrants may face lower legal opportunities, meaning that they have on average lower $w'(\tau^l)$, or higher μ (see, for example, Algan at al, 2010). For instance, they hold lower quality jobs and a lower chance of getting accepted in higher status jobs. This might be because of discrimination, or limitations in language proficiency. Thus, they may find opportunities in illegal sectors more attractive.

On the other hand, there are a few reasons which would indicate a negative association between being an immigrant and criminal behavior. Some evidence by criminologist shows that the criminal justice system and law enforcement are biased in various stages against ethnic minorities (see for example, Smith, 1997, Feilzer and Hood, 2004). This implies that immigrants may face more severe punishments compared to natives. Moreover, highly deprived areas are generally associated with both higher concentration of immigrant population and higher concentration of police force. This increases the risk of apprehension. Finally, immigrants also face deportation which is a punishment specific to them, which could be considered as a large disincentive to commit crimes (Butcher and Piehl, 2007). Thus, according to the above, we would expect that the average immigrant faces both higher $\pi(\tau^i)$ and $P(\tau^i)$.

In another direction, discount factors and risk attitudes may also be different for immigrants. It could be said that immigrants are willing to take more risks, since migration is in general a risky action with quite uncertain outcomes (see, for example, Jaeger at al, 2010). On the other hand, there is empirical evidence that immigrants are more risk averse than the native population (see, for example, Bonin at al, 2009). A considerable number of immigrants leave their families back on their source countries. Even though they have taken the risk to migrate away from their countries, they target on a better life for them and their families. In addition, they would like to feel socially equal to natives by presenting a highly responsible and credible behavior. They may not be willing to take highly risky actions which can cost them their presence in the host country. In addition, a large proportion of immigrants come from poor countries. Since they have already faced quite harsh conditions, it could be assumed that they are more resilient not only in financial difficulties but also in psychic and physical severities. These are only a few examples to explain why we would expect some differences between risk attitudes of natives and immigrants.

Furthermore, discount factors and risk behavior are strongly associated with cultural factors. Therefore, coming from different cultures, risk attitudes and discount factors may be shaped quite differently. Cultural differences are also important for the perceptions towards the moral dimension of crime. Thus, psychic costs, also incorporated in $k'(\tau^i)$, may be very different between immigrants and natives.

Finally note that, the model does not explicitly include variables for demographic factors such as age, gender, or location features, that are found to be associated with crime. Therefore, there could be also some indirect effects of immigration on crime if immigrants are different from natives with respect to these demographic features.[23] Thus, taking all the above discussion into consideration, the individual relationship between immigration and property crime is not determined by theory, and can only be established by an empirical analysis, using a well specified model and appropriate data.

3 Immigration and Crime. A Review of Research

Although other indicators of crime, such as education, inequality, labor opportunities, etc., have been well studied by economists, the empirical research of the immigration-crime nexus is limited. However, the literature by criminology and sociology scholars is much more extended, both theoretically and empirically. Traditionally, these studies are developed in countries which have experienced large inflows of migrants. For instance, the US with the inflow of Latino and Afro-Caribbean population, Germany with immigrants from Turkey and the former Yugoslavia, Netherlands with Turkish and Moroccans, etc.

Several times, the results of various researchers are contradictory. This is natural, mainly because the empirical results by each researcher are subject to the composition

[23] All the discussion above concerns the individual supply of property crime that economic theory predicts. In another direction, long before economists, criminologists developed some ideas on the immigration-crime nexus. Starting with eccentric ideas that immigrants commit more crimes just because they are a group of inferior individuals (see, Armstrong, 1935, and, Sellin, 1938), they switched to more rational theories based on psychological patterns. One of the earliest theories is based on the so-called "strain" theory, presented by Merton (1938), which states that immigrants present adverse behavior due to accumulative pressure, as for example, because of discrimination, racism and unequal social and financial opportunities. Other theories suggest that there might be deviant behavior by both immigrants and natives because of cultural conflicts. Thus, contrary to economic theory, criminologist's theories stand better for violent crime. For an excellent survey on these theories the reader may refer to Martinez and Lee (2000).

of immigrant population in each destination country, the circumstances that immigrants encounter in different countries, the differences in the data sets they use, and last but not least, the different statistical tools and strategies each researcher follows. Thus, we cannot identify globally what is the effect of immigration on crime by looking only at one country, or one approach, but we need to look at the broader picture.

The literature review is presented in two subsections. In the first one the results found by economicsts are presented, whereas the second one briefly presents the results found by sociologists and criminologists. A basic difference between studies by criminologists and sociologists and studies by economists is the theoretical hypothesis. The first group bases its theory on disorganization and culture conflicts whereas economists associate immigrants with crime through the economic models of crime. The second main difference stems from the fact that economists traditionally use more analytical statistical and mathematical tools than the other group. Thus, in general the studies by economists use more sophisticated and in many times, more appropriate statistical models.[24]

3.1 Empirical Evidence by Economists

To begin with, as mentioned in the introduction, to my knowledge there is only one study concerning the UK. Bell, Machin and Fasani (2010) examine how two separate large waves of immigrants affected crime rates. These waves are, the late 1990s wave of asylum seekers and the large inflow from the "A8" Eastern European countries since May 2004. What they find is that the first wave is associated with higher property crime, even after controlling for endogenous location using fixed effects and instrumental variables.[25] However, they find that the A8 wave did not affect property crime.[26] Moreover, their results indicate that there is no effect for violent crime. They argue that this finding is consistent with a simple economic model of crime, as asylum seekers face much lower legal opportunities relative to A8 immigrants and natives, and therefore, illegal activities seem more attractive to them.

[24] This review does not intend to criticize or judge the methods and specifications of different researchers, but instead, it is purely descriptive.

[25] As asylum seekers where located by the National Asylum Support Service, they were mostly located in unpopular areas with a large amount of vacant houses. Thus, they instrument for endogenous location decisions by the number of dispersal accommodation in each local area.

[26] In this case they control for endogenous location by using the availability of flights to A8 countries as an exogenous variation for immigrants choices of location.

As far as I know, the first attempt by economists to investigate the immigration-crime relationship is that of Bucher and Piehl (1998a). Using data from the Uniform Crime Reports and Current Population Surveys, they first look at the aggregate effects of immigration on crime in the US, during the 80s. Although they find that there is a positive relationship between crime rates and the fraction of recent (within one year) immigrants, this association fades out once they include controls both correlated with the location choice of immigrants and crime rates. Actually, the effect of immigration becomes negative but statistically insignificant. Using fixed effect analysis, they find that there is no association (negative but insignificant) between flows of immigrants and crime rates, or flows of immigrants and growth in crime rates (one year changes). Their results are strengthened by the use of self reports from the National Longitudinal Survey of Youth of 1980. Their sample includes respondents between 15 and 23 years old, and they also separate their sample between men (5,205 people) and women (5,504 people). What they find is that in all cases immigrants report considerably less crime (statistically significant at 5% in most cases), result that it is reinforced once they control for other individual characteristics associated with crime. Immigrants are also less involved in arrests and convictions. Nevertheless, they do not use any strategy to control for any possible under-reporting, a major concern in self-reports.

Bianchi, Buonanno, and Pinotti (2007), study the same research question but for Italy, by using police administrative records for Italian provinces. Using a panel data set from 1990 to 2003, they find a positive relationship between the size of immigrant population and most categories of crime rates, even after controlling not only for other determinants correlated with the factors that determine both crime and the location choice of immigrants, but also for province and year dummies. However, they recognize that even after controlling for these factors, there can still be some time varying unobserved factors correlated with both immigration and crime (for example, a economic crisis in a specific area would reduce the cost of living, which would in turn attract immigration population, but on the other hand, also increase crime rates). Moreover, there is the concern of reverse causality, since crime rates in an area could affect the location choice of immigrants. Therefore, they employ a two-stage least square approach, using changes over time of immigrant population in the rest of Europe as an instrument for changes of immigrant population in Italy. Arguing that this is not a "weak" instrument, they find that there is no relationship between immigration and most categories of crime. However, there still exists a positive association for murders, robberies and thefts.

Finally, in another direction Bucher and Piehl (1998b, 2007), use institutionalization rates as a proxy for incarceration rates. Using the 5% Public Use Microsamples of the US census in 1980, 1990, and 2000, they find that the probability of an immigrant being incarcerated is much lower, even after controlling for educational attainment and ethnic status. They also find that this difference has increased over the last three decades, as more recent immigrants have the lowest incarceration rates. They attribute this to two reasons. First, they argue that this improvement is due to the stricter legislation for immigrants, since recent laws have broadened the crimes for which an immigrant is deported. They find that although deportation itself does not drive the results (meaning that the share of immigrants is not lower in prisons just because they get deported), it acts as a deterrent effect specific to immigration population. Second, they show evidence that the recent migration process to the US selects individuals who are either less crime-prone, or more responsive to deterrent effects.

3.2 Empirical Evidence by other Scholars

On the other hand, criminologists and sociologists have paid more attention to this relationship. Here, I will try to briefly describe the results of the most important works in each country.

The majority of immigration-crime studies come from the US and focus on violent crime.[27] Most of evidence from the US shows that although the public opinion keeps associating immigrants with crime, immigration is not associated with higher crime, and in many cases it is association with lower crime. Hagan and Palloni (1998), perform an empirical analysis using a sample of 34 Standard Metropolitan Statistical Areas of the

[27] A more detailed review can be found in Hagan and Paloni (1998), and Martinez and Lee (2004). For most recent evidence refer to Stowell (2007). Also, each individual study described in the subsection provide some related literature. For example see, Ousey and Kubrin (2009), Stowell at al (2009), and Wadsworth (2010). Stawell at al (2009), also provide a table that presents important information for the main studies in the US (Table 1, p.895).

US. By regressing logged arrests on the proportion of immigrants in the population find no association between immigration and both property and violent crime. Reid at al. (2005), combining 2000 US Census data and 2000 Uniform Crime Report (UCR), explore how the immigrant population affects crime rates across a sample of metropolitan areas. They find that, after controlling for demographic and economic conditions, immigration does not affect violent and nonviolent crime. Lee, Martinez and Rosefend (2007), examine whether immigration increased homicide in the three border cities of Miami, El Paso, and San Diego, using 3,345 homicide occurrences happened between 1985 and 1995. Poisson regression results indicate that there is no relationship, or, even a negative relationship between the percentage of new immigrants and homicide levels. Stowell (2007), in the same direction examines the crime-immigration nexus for three US cities, Alexandria, Houston and Miami. Using neighborhood-level data from 2000, he also finds no direct evidence between the proportion of recent (less than ten years) immigrants in the population and violent crime levels. However, he finds a negative association for Miami.

Even more recently, there was a quite large amount of publications on this subject for the US.[28] Most of those studies also show that if there is an association, this is negative. Very briefly, Ousey and Kubrin (2009), using fixed effects for 159 US cities (with more than 100,000 residents) for the three time periods 1980, 1990 and 2000, find that the proportion of recent immigrants decreases violent crime. Wadsworth (2010), in a similar manner, uses a fixed effect model for 459 cities (with more than 50,000 residents), between 1990 and 2000. Both robbery and homicide rates are examined. The results of this study suggest that the proportion of foreign born population and the proportion of new immigrants decreased crime rates within this time period. Stowell at al (2009) use a panel over the period 1994-2004 for 103 metropolitan areas (with more than 500,000 residents). Their results indicate that there is a negative association between changes in immigration concentration and changes in crime rates. This effect is particularly stronger for robberies.

The "Homicide Studies" journal published an issue on 2009 that solely focuses on aspects between immigration and homicides and other crime types. Graif and Sampson (2009), to a neighborhood study of Chicago, using a "weighting" estimator that assigns different

[28]For all studies described in this paragraph, crime data are collected from the Uniform Crime Reports.

weights for points of different proximity to each data point, find that higher concentration of foreign born population is either negatively or not associated with homicide. Feldmeyer and Steffebsmeier (2009), using homicide arrest data from California, find that the proportion of recent (entered USA between 1990-2000) immigrants does not affect the mean number of overall homicides, but it does affect negatively the mean of homicides against white and black people. Vélez (2010), by allowing the effect of immigration to be different between advantaged and disadvantaged neighborhoods of Chicago (by using an interaction term), finds that an increase of recent immigrants decreases the number of homicides in disadvantaged areas but has no effect in the more advantaged ones. In another study for Chicago, Chavez and Griffiths (2009), argue that growing immigrant population was unrelated to homicide patterns. Polczynski at al (2009), look at arrest rates for different types of crime in Orange County, Florida, and show that arrest rates are generally lower for foreign born individuals. Moreover, their results suggest that concentration of immigration is not associated with the number of arrests. Finally, the studies of Akins, Rumbaut, and Stansfield (2009), and Stowell and Martinez (2009), focus on the area of Miami. The former suggests that for the area of Austin, Miami, where the immigration population increased by around 580 percent, there is no association between migration and homicides. In the latter, they find that neighborhoods with higher number of Latino immigrants exhibit lower levels of homicide.

In Germany the studies of immigration-crime link are based on official criminal statistics, since in official records from police or courts there is a categorization of people as foreigners or not. Using police and court data Albrecht (1987) finds that foreign population's involvement in criminal activities is higher than that of Germans. However, this relationship disappears once controls for socio-economic conditions and demographic differences are accounted for. The same positive relationship is also found in Albrecht (1997). However, in this second study the higher involvement of immigrants in crime persists, even after controlling for the above factors. Finally, Chapin (1997), using basic statistic analysis finds that changes in foreign population increase the growth of crime levels.

Evidence from official statistics of Switzerland also suggests that immigrants are less law-abiding. For example, Killias (1997) using police and conviction statistics finds that, although in the 70s immigrants displayed similar crime rates with natives, after the 80s immigrants were over-represented in crime statistical tables. However, he expressed many concerns about the reliability of official crime statistics. Contrary to that, using self-reported data from more than 3,000 adolescents and employing basic statistics, Vazsonyi and Killias (2001), find that first generation adolescent immigrants display slightly lower crime rates than native Swiss adolescents. This result coincides with other works in Switzerland that have used self-reported crime data. They also find that second-generation immigrants are more crime-prone than natives, a result that it is common in the literature.

In Netherlands, Junger-Tas (1997), finds that Moroccans and Antilleans are over-represented in official criminal statistics. Contrary to that, he also presents a review of self-report studies in Netherlands which suggest that the above groups, and other groups of ethnic minorities, are less involved in crime. Concerning France, Tournier (1997), finds that although immigrants are over-represented in prison statistics and statistics of criminal suspects, a large fraction of their crimes concerns violations of immigration law regulation. He shows that if these violations are taken into account, the difference between foreigners' and natives' involvement in crime would be considerably lower. In a study for Sweden, Martens (1997), finds that the fraction of immigrants who have been suspected for crimes relative to their population is clearly higher than that of native population. A difference still exists even after immigrant-native differences are accounted for, although it becomes noticeably smaller. He also finds that first-generation immigrants display higher crime rates than second-generation immigrants, a result that contradicts with findings in other countries. Finally, Yeager (1997) presents a cross-country review of immigration and criminality. According to this review, immigrants are not as highly involved in crime as natives in Canada and Australia. Criminal records form France, Sweden, Netherlands and Germany also indicate that immigrants are over-represented in various criminal aspects. He also briefly describe the results of crime self-reports for Switzerland, Netherlands and Germany, which suggest that immigrants are less involved in crime than natives citizens.

4 The OCJS. Some Methodological Issues

The basic target of the present study is to identify whether immigrants are more or less crime-prone than natives, even after controlling for the fact that immigrants might exhibit some differences in basic demographic characteristics associated with higher or lower crime. For this purpose the Offending Crime and Justice Survey (OCJS) of 2003 is used, a nationally representative survey which asks people in England and Wales about their experiences as offenders and their attitudes towards criminal activities.[29] Although a few earlier offending surveys in the UK exist, this is the largest one and the most sophisticated in terms of design and construction.[30] This section discusses some features, advantages and limitations of the OCJS, and self-report studies in general.

A basic target of the OJCS is to precisely measure the prevalence of offending in the general population, as opposed to studies of already convicted criminals, and to investigate the factors related with these crimes (Hamlyn, and Hales 2003).[31] However, this is a hard task if we consider a few limitations of the OCJS. To begin with, validity and reliability of the responses is a concern, since questions try to elicit information in a very sensitive part of personal activities such as crime. Particularly, a high response error is expected that most probably takes the form of under-reporting. However, we need to stress that computer-based interviews are used as opposed to face-to-face interviews, which increases the reliability of the responses (see, Turner at al, 1998). As will be explained in Section 5, a conventional regression model that does not take into acount under-reporting would result in inconsistent estimates for the determinants of the true crime, as it is developed to estimate the coefficients of the observed, under-reported, crime. More importantly, if immigrants under-report more than natives, the estimated effect of being an immigrant on crime will be downward biased.

Another concern stems from the fact that some selected by the survey's conductors did not accept to participate on the survey. In spite of the fact that response rates of

[29] Although three subsequent OCJSs exist (2004, 2005, 2006), they are particularized in adolescent delinquency (people from 10 to 25 years old). Thus, they are not appropriate for the purposes of my research question.

[30] For a review of other self-report studies the reader can refer to Farrington (2003), Thornberry and Krohn (2000), and Junget-Tas and Marsall (1999).

[31] For surveys of convicted criminals see, Budd at al. (2005).

the OCJS are very close to response rates of other population surveys, such as the Labor Force Survey or the General Household Survey (see, Sharp and Budd, 2005), estimates of prevalence of crime will be downward biased in case non-respondents commit more crimes (see, Farrington at al., 1990).[32] In addition, if the effect of being an immigrant on criminal behavior is different between respondents and non-respondents, the coefficient measuring the difference between immigrants and natives propensity to crime will be biased. Moreover, this survey does not capture institutionalized individuals who would most probably commit more crimes than the general population if they where free. The consequences are similar to the previous point.

Despite these limitations, relying on self-reports constitutes the most suitable method to identify predictors of criminal behavior. As Thornberry and Krohn (2000) point out, the best way to identify factors of criminal behavior would be to observe the actual behavior of potential criminals, self reports being the nearest proxy to actual criminal behavior. Particularly, the OCJS also provides important information on a wide range of characteristics of potential criminals as opposed to victimization surveys and official statistics.[33] Moreover, it would be misleading to attempt to identify the determinants of criminal behavior by comparing convicted and non-convicted individuals, since there is quite a large number of individuals that have committed crimes but are not convicted. Similar logic applies to comparisons between arrested immigrants and natives.

As we saw in the previous section, some criminologists prefer to use data of recorded crime. There are two main pitfalls in using official recorded crime by the police. First of all, these statistics are very poor in terms of information provided. Concerning official statistics in England and Wales, they offer no information concerning immigration status. Therefore, they are inappropriate for native-immigrant crime comparisons. Furthermore, even if the same information was available, it is widely accepted that many crimes are not reported to the police, and many reported crimes are not even recorded because of inside-police operational reasons (for example, some reported crimes are not considered

[32] Although the weights used in empirical analysis take into account non-respondents (see, Hamlyn and Maxwell, 2003, and Budd, Sharp and Mayhew, 2005), they do not control for possible higher crime of non-respondents.

[33] Nowadays, self-reports is the most commonly used technique among researchers to discuss causes of crime. (See, Hagan, 1993, and Junget-Tas and Marsall, 1999).

serious enough to be recorded). This is the so-called "dark-figure" of crime, for which a vast literature exists (see for example, MacDonald, 2001, 2002). Moreover, there is some evidence by criminologist that the criminal justice system and law enforcement are biased in various stages against ethnic minorities, particularly against black people (see for example, Smith, 1997, Feilzer and Hood, 2004). Perhaps, immigrants are also more "visible" by the police because of over-policing in target areas where ethnic minorities are concentrated, increasing the chances of immigrants to be arrested (Sharp and Budd, 2005). Therefore, official statistics would overestimate immigrants' crime if immigrants face a higher probability of arrest for the same crimes, and if police officers disproportionately record crimes that are supposedly committed by immigrants.

On the other hand, although victimization surveys, such as the British Crime Survey (BCS) for England and Wales, provide the most precise estimates of the actual crime, they do not provide any information about criminals' characteristics. However, if we accept that the BCS reveals the figure of crime that is the closest to the true one, we could compare official and OCJS' crime figure to BCS' in order to evaluate their precision in measuring crime. Kershaw and Walker (2008), suggest that recorded crime is only 42% of the total crime in England and Wales. Concerning the OCJS, Budd, Sharp and Mayhew (2005), suggest that the figure of violent crime is quite close to that of BCS, but the count of property crime is quite lower than in BCS. This would suggest that there is under-reporting in property crime. However, these facts must be treated with caution since there are fundamental design differences between these two sources, and therefore, they do not provide comparable crime figures (see, Budd, Sharp and Mayhew, 2005).

5 Econometric Models

As will be described in the next section, the dependent variable is observed in count form (number of crimes committed in the last twelve months), therefore, both count data models and binary choice models can be specified. However, as will be better explained in the next section, the very large number of zeros in the property crime variable, resulting in very low variation in the dependent variable, will make estimation of these models quite harsh, mainly when estimators to allow for under-reporting are used. Alternatively, a safer choice would be to use the binary information, whether or not someone has committed a crime last year although these models do not use all available information. Therefore, a compromise would be to base my empirical results on binary choice models, while count data estimation models can be used for robustness check analysis.

As explained in detail before, under-reporting is the main concern in this data set.[34] In both cases of binary and count data models, it has been shown that the resulting estimates from traditional estimation techniques are inconsistent if under-reporting, or more generally, misreporting is present (Hausman at al, 1998, Cameron and Trivedi, 1998 (p.313), and Papadopoulos, 2010). On this direction, parametric models that take into account misreporting (both under and over-reporting for binary choice models, but only under-reporting in count data models) will be used. Another concern is the low variability of the dependent variables as there is a quite large number of zeros (95%). This has obvious consequences for the precision of the estimates in both binary and count data models. Adding the fact that estimation of models that account for under-reporting is quite demanding, to achieve precise estimates quite rich samples are required.

Moreover, as explained in the previous section, there is also a sample selection problem, since it is likely that people who refused participating in the survey are more prone to crime. However, models that correct for sample selection problems, such as the Heckit procedure developed in Heckman (1976), would require information of not respondents, which is not available. Therefore, this problem is ignored in the analysis, hoping that non respondents exhibit the same criminal behavior, or at least that the crime differentials between immigrant and natives non respondents follow the same pattern with crime differentials between immigrant and native respondents.

In the following subsections models that control for misclassification, or, under-reporting are described.

[34] Although we could assume that someone would never report a crime if he has not committed one, we could not rule out that over-reporting may be present as well. This could be attributed to the fact that the OJCS is a retrospective study and measurement error in both directions could be possible. Although over-reporting is most possible unintentional, so that it is random, under-reporting is most probably intentional.

5.1 Binary Choice Models

In this section the model of Hausman at al (1996, 1998) is presented, a parametric model that takes into account both probabilities of under-reporting (misclassification of a true one as zero) and over-reporting (misclassification of a true zero as one).

The model comes naturally from a latent variable specification. To simplify things, assume that in a given period of time, an individual would commit a crime (or a number of crimes) if the total utility from committing this crime is higher than the utility obtained from not committing it. So, let u_i^* be the (unobserved) utility obtained if committing the crime(s) minus the utility if not committing it (them). We specify u_i^* to depend linearly on the vector of characteristics x_i such that,

$$u_i^* = x_i'\beta + \epsilon_i, \tag{5.1}$$

where ϵ_i is a random error. Thus, the individual commits at least one crime according to the following,

$$y_i^* = \begin{cases} 1 & \text{if } u_i^* > 0 \\ 0 & \text{if } u_i^* \leq 0 \end{cases} \tag{5.2}$$

where y^* is a dummy for the true but unobserved crime. According to (5.2), conditional on x_i', the probability of committing a crime is given by,

$$\Pr(y_i^* = 1|x_i) = \Pr(u_i^* > 0|x_i) = \Pr(\epsilon_i > -x_i'\beta|x_i) = F(x_i'\beta) \tag{5.3}$$

where $F(x_i'\beta)$ is assumed to have a known functional form such as the standard normal.

Let us now define y_i to be a dummy for the reported and therefore, observed crime. Suppose that the reported crime does not coincide with the actual crime since there is mis-reporting (in this context defined as misclassification). The probabilities of misclassification are defined as follows,

$$
\begin{aligned}
a_1 &= \Pr(y_i = 1|y_i^* = 0), \\
a_0 &= \Pr(y_i = 0|y_i^* = 1),
\end{aligned}
\tag{5.4}
$$

where a_1 is the probability of reporting a crime, conditional on committing no crime (over-reporting) and a_0 is the probability of reporting no crime, conditional on committing crimes (under-reporting).

It is easy to derive the conditional probabilities of the observed crime, incorporating the probabilities of misclassification. The response tree in figure 5.1 is a very clear way to do that. It is clear that these probabilities are given by,

$$
\begin{aligned}
\Pr(y_i = 1|x_i, a_0, a_1) &= (1 - F(x_i'\beta))\, a_1 + F(x_i'\beta)(1 - a_0) \\
&= a_1 + (1 - a_0 - a_1)F(x_i'\beta)
\end{aligned}
\tag{5.5}
$$

$$
\begin{aligned}
\Pr(y_i = 0|x_i, a_0, a_1) &= (1 - F(x_i'\beta))\, (1 - a_1) + F(x_i'\beta)a_0 \\
&= 1 - a_1 + (1 - a_0 - a_1)F(x_i'\beta)
\end{aligned}
\tag{5.6}
$$

We can estimate, a_0, a_1, and, β, using the method of Maximum Likelihood (MLE) once we have specified the log-likelihood function as,

$$
\ln L(\beta, a_0, a_1) = \sum_{i=1}^{n} \{y_i \ln [\Pr(y_i = 1|x_i)] + (1 - y_i) \ln [\Pr(y_i = 0|x_i)]\} \tag{5.7}
$$

Given correct specification of the model, meaning that the specified model of constant misclassification is the correct model under the true data generating process (dgp), maximization of (5.7) using numerical optimizers, such as the Newton-Raphson, yields consistent estimates for the coefficients of true crime, the probability of under-reporting, and the probability of over-reporting.

We notice that a_0 is only designed to capture total under-reporting as opposed to partial under-reporting. That is, the probability of under-reporting cannot capture the cases where individuals report a portion of the total number of crimes they have committed. However, as is explained later in this section, models that use the count form of the dependent variable are able to estimate the probability of any particular crime to be reported.

Hausman at al (1998) show that the model is not globally identified since, $a_1 + (1 - a_0 - a_1)\Phi(x_i'\beta) = \widetilde{a_1} + (1 - \widetilde{a_0} - \widetilde{a_1})\Phi(-x_i'\beta)$ where $\widetilde{a_0} = 1 - a_1$ and $\widetilde{a_1} = 1 - a_0$. Thus, there are two observationally equivalent models with parameters (a_0, a_1, β) and $(\widetilde{a_0}, \widetilde{a_1}, -\beta)$. Identification is achieved by imposing the "monotonicity" condition, which states that $a_0 + a_1 < 1$. According to this, we are able to rule out the "wrong" maximum, since $a_0 + a_1 < 1$ implies that $\widetilde{a_0} + \widetilde{a_1} > 1$. If this condition fails, the misclassification probabilities are too large, and therefore, the data are most probably too noisy to obtain reasonable results.

It is clear from (5.4) that the probabilities of misclassification do not depend on individual characteristics. As this is not realistic in many application, including the present study, this assumption can be easily relaxed if we model a_0 and a_1 to be conditional on a set of covariates, so that, $a_0 = F(z_i'\gamma)$, and $a_1 = F(w_i'\delta)$, where F can be the cumulative distribution of any binary model. The vectors of the regressors (x, z, w) can be the same, disjoint or overlapping. No further assumptions are required to identify this model.

Although this is quite an easy model to implement in econometric software, in practice estimation is quite difficult, particularly when the misclassification probabilities are allowed to depend on variables, and when the data are quite noisy with very low variation particularly in the dependent variable. Always, exclusion restrictions could help the estimation procedure. Moreover, as Hausman at al (1996) note, it is quite difficult to get precise estimates, and it seems that quite rich samples are needed. There are some papers in the literature that have utilized this estimator (see for example, Leece, 2000, Artis at al, 2002, Caudil and Mixon, 2005, and Falaris, 2007). Nevertheless, none of them allows the probabilities of misclassification to depend on covariates.[35]

So far, we have regarded zeros coming from two different sources; misclassification of 1

[35] It is interesting that this estimator is very similar to the Detection Controlled Estimator, presented in Feinstein (1989). In that paper Feinstein explains that sometimes inspectors fail to detect a violation. However, they never detect a violation if there is not one. Therefore, in the simplest form, when the probability of detection and the probability of violations are independent, he derives the same log-likelihood function as (5.7), if we set a_1 to zero. However, in his estimator, a_0 depends on regressors. This concept can be naturally applied in any situation that involves compliance and inspection, as for example, tax evasion (Feinstein, 1991).

as 0, and zeros from the traditionally binary choice model. Nonetheless, zero-inflation can be incorporated into this model if we think of zeroes coming from a third source. That is, there are some individuals who, regardless of the conditioning set x_i, never commit any crimes (and consequently they do not report any).[36] If we do not incorporate this zero-inflation probability separately, the estimated probability a_0 cannot distinguish under-reporting from zero-inflation and as a result, it must be regarded as a mix of the probability to under-report and the probability of never committing any crimes.[37] This is clear if we examine the second part of figure 5.1, where a_0 and a_1 are expressed as probabilities of zero-inflation and one-inflation respectively. Following this tree we notice that the conditional probabilities of observing 1 and 0 are exactly as in (5.5) and (5.6). However, in this case the interpretation of a_0 and a_1 is very different. Although, probability of one-inflation does not make a lot of sense, zero-inflation is quite possible.[38] Nevertheless, it is possible to separate under-reporting from zero-inflation by incorporating zero-inflation separately in the likelihood function. This model is presented in Appendix A.

5.2 Count Data Models

The models presented here are based on the Poisson Logistic Regression model of Winkelmann and Zimmermann (1993), also presented in Mukhopadhyay and Trivedi (1995).[39] This model is a particular case of a Poisson-Stopped Sum distribution where the iid random variables to be summed follow the Bernoulli distribution with constant probability of

[36] Think of them as genuinely non criminals.

[37] This idea of adding in a probability of inflation can be traced back to Gaudry and Dagenais (1979) where they developed an estimator for multinomial choice models, calling this the "dogit" model. Gaudry (1980) and Swait and Ben-Akiva (1987), apply this model for individual choices between a set of different transportation modes. They explain that given a set of choices, an individual is either captive to one choice regardless of his/her characteristics (inflation probability) or free to choose from the full set of choices (traditional multinomial choice model). In this models there is inflation probability for each category, whereas in the model presented here there is only zero-inflation probability, making this model more similar to zero-inflation models for count data as described in Mullahy (1986) and Lambert (1992).

[38] One-inflation could be interpreted as follows: There are some individuals who, regardless of their characteristics, always commit crimes in the specified time period, and they are always willing to report that they have committed them.

[39] Here, only a brief discussion of these models is presented. For a more detailed analysis the reader may refer to Winkelmann, 2008, Cameron and Trivedi, 1998, and Papadopoulos, 2010a.

success p. According to this specification, the observed number of counts is given as,

$$y = b_1 + b_2 + \ldots\ldots + b_{y^*}, \tag{5.8}$$

where y^* is a latent variable of the true counts that follows the Poisson distribution with parameter λ.[40] A basic assumption of any Stopped Sum distribution, is that y^* and b_i are conditionally independent.[41] Under this independency condition it is easy to show (for example, using probability generating functions as in, Feller, 1968) that the observed counts, y, also follow the Poisson distribution with parameter $\mu = p\lambda$. Of course, it is clear that $y \leq y^*$, and it is said that the observed counts are "under-reported". We must underline that this specification assumes no over-reporting.[42]

The above concept can be extended in a multiple regression framework, where the probability of success and the true Poisson process are allowed to depend on covariates. In the Poisson-Logit model, the true counts follow the Poisson distribution with,

$$\begin{aligned} &\Pr(Y_i^* = y_i^* | x_{1i}) = e^{-\lambda_i} \lambda_i^{y_i^*} / y_i^*!, \\ &\lambda_i = E[y_i^* | x_{1i}] = e^{x'_{1i}\beta}, \end{aligned} \tag{5.9}$$

and the probability of success is given as a Logit, so that,

$$\Pr(b_{ij} = 1 | x_i) = \Lambda(x'_{2i}\gamma) = \Lambda_i = \frac{e^{x'_{2i}\gamma}}{1 + e^{x'_{2i}\gamma}}. \tag{5.10}$$

Consequently, it can be shown that the observed counts also follow the Poisson distribution (see, Papadopoulos, 2010a) with conditional probability distribution

$$\begin{aligned} &\Pr(Y_i = y_i | x_{1i}) = e^{-\mu_i} \mu_i^{y_i} / y_i!, \\ &\mu_i = E[y_i | x'_{1i}\beta] = \lambda_i \Lambda_i, \end{aligned} \tag{5.11}$$

[40] The name Stopped-Sum comes from the fact that the summation of Bernoulli variables is *stopped* by the value of the Poisson distributed latent variable y^*.

[41] This assumption is relaxed on Winkelmann (1998).

[42] At least we need that the measurement error, over-reporting, is not correlated with the regressors, so that it is totally random.

and log-likelihood function given by,

$$\ln L(\beta,\gamma)=\sum_{i=1}^{n}\left\{-\mu_i+y_i\ln\mu_i-\ln\left(y_i!\right)\right\}. \tag{5.12}$$

Papadopoulos and Santos Silva (2008), show that identification of this model requires at least one exclusion restriction in the Poisson process, or at least one sign restriction on a parameter of the Logit part (see also, Section 7).

Despite the fact that consistency of the Poisson-Logit estimator only requires that $\mu_i = E[y_i|x_{1i}]$, we can extend this model to account for possible over-dispersion, relaxing the strong assumption that both mean and variance of y_i are equal to μ_i. The standard way to do this is by adding an unobservable individual effect ϵ_i which will account for extra unobserved heterogeneity. Now, conditionally on x_i and ϵ_i, y_i has a Poisson distribution with parameter $\mu_i e^{\epsilon_i}$. Under the usual assumption that e^{ϵ_i} has a gamma distribution with unit mean and variance α_i, the distribution of y_i, after integrating out the unobservable individual effect, is negative-binomial with mean μ_i and variance $\omega_i = \mu_i + \alpha_i\mu_i^2$. If the variance of ϵ_i is constant (homoscedastic), we obtain the Negative Binomial 2-Logit (NB2-Logit). As Papadopoulos and Santos Silva (2008) show, the conditions required for identification of this model are exactly the same with the Poisson-Logit model (see also, Section 7, for more details).

The log-likelihood of the NegBin2-Logit is the following,

$$\begin{aligned}\ln L(\alpha,\beta,\gamma)=\sum_{i=1}^{n}&\ln\left(\Gamma(y_i+\alpha^{-1})/\Gamma(y_i+1)\Gamma(\alpha^{-1})\right)\\&-(\alpha^{-1}+y_i)\ln(1+\alpha\mu_i)+y_i(\ln\mu_i+\ln\alpha).\end{aligned} \tag{5.13}$$

Maximization of (5.16) will yield consistent estimates for (α,β,γ), given correct specification of the model, that is, the true dgp is NegBin2-Logit. To provide evidence against equidispersion a lagrange multiplier test can be performed since Poisson-Logit and NegBin2-Logit are nested, as can be shown that NegBin2-Logit reduces to the Poisson-Logit when α is zero. Other possible NB-Logit models considering a different functional form for the variance are presented in Appendix B. Finally, we need to stress that similarly to the models for binary choice, models for count data with under-reporting can also be generalized to take into account zero-inflation. These models are also presented Appendix B.

6 Data Set and Discussion of Variables

Before describing the details of the data set, it must be stressed that some effort has been made to keep the sample size as large as possible. The reason behind this is twofold. First, the econometric methods used in the empirical analysis are very demanding, as explained in the previous section. Secondly, the variation of the dependent variable is very low, as 95% of respondents reported no property crime. Therefore, larger samples will assist on identifying the coefficients of interest more easily by obtaining more precise estimates.

To achieve the highest sample size possible, I exploit the sophisticated design of the OCJS that makes use of "boost" samples. There are three independent samples in the OJCS; the core sample (10-65 year olds, 58%), the young boost sample (10-25 year olds, 27%), and the non-white boost sample (non-white individuals, 10-65 years old, 15%).[43] Each sample is accompanied by its (sampling) weighting variable, which must be used to restore the representative of each sample. I increase the sample size by around 5000 individuals by adding these three samples together. However, to re-establish representativeness, a weighting variable that combines the three separate weights is used.[44] A tabulation by sample type follows in Table 6.1. Thus, the resulting data set I use in my empirical investigation consists of 11,658 individuals, 5,604 males and 6,054 females, between 10 and 65 years old.[45]

[43]Criminal behavior of people between 10 and 25 years old is the primary interest of the OJCS. The subsequent OJCSs of 2004, 2005 and 2006 include only 10-25 year olds individuals. Some of them are included in all OJCSs constituting a panel. An ethnic boost sample has been also constructed for investigating crime for this group of people separately. The percentages in the parentheses denote the fraction of people of each sub-sample to the total sample size. This is the total sample size I use in my analysis and not the initial total sample size of the unrefined OJCS.

[44]This weighting variable was kindly provided by the Home Office. A detailed analysis of the construction of the combined weighting variable is given in the Appendix of Hamlyn and Hales (2003).

[45]We need to note that by using individuals from 10 to 65 years old, we include students, retirees and people not in the labor force such as house keepers. Therefore, there is a departure from the economic model of crime which is designed for people in the labor force. However, these groups could somehow fit in the model as we can think that students are looking at their future legal opportunities, housekeepers are considering the household income, and retirees receive a legal stream of pensions. Anyhow, in the empirical analysis we try to formulate and estimate the behavior towards property crime for the whole population. This was inevitable, as limiting the sample only to individuals in the labor force reduces the sample size at a point where obtaining sensible results seems impossible (at least for the models with under-reporting

Table 1: Tabulation by Sample Type

Sample Type	N	%
Core 10-65	6,771	58.08
Boost 10-25	3,098	26.57
Ethnic Boost	1,789	15.35
Total	11,658	100.00

Concerning the dependent variable, as explained before, this paper focuses on property crime.[46] The information on property crime can be either used to construct a dichotomous variable which takes the value one if someone has committed a property crime last year, or a count variable measuring the number of property crimes in the year prior the interview. The latter will be used for robustness analysis only, as the large number of zeroes complicates the estimation of count data models.[47] The observed distribution of the property crime variable is given in Table 6.2. We observe that 94% is concentrated in 0 crimes, 5.5% is concentrated between 1 and 14 crimes, and the rest 0.5% is spread from 15 up to 225 crimes. Using the sampling weights, we find that the unconditional mean is 0.34 crimes per person, whereas the unconditional variance is 19.44. Therefore, according to the raw data there is strong evidence against equi-dispersion, which suggest that Negative Binomial models may provide a better fit to the data. Moreover, notice that the sample size differs between the count and the binary form of the variable. This is because some of the people who reported a crime were reluctant to report the number of crimes they committed. Consequently, these observations are recorded as missing cases, leading in further reduction in the variation of the dependent variable.

Information about violent crime is also available.[48] Although violent crime is not the main subject of this paper, a few empirical results of violent crime are presented in subsection 7.9. Moreover, it would be possible to separate crime in more crime types (such

and misclassification).

[46]Property crime consists of: thefts and attempted vehicle thefts, thefts and attempted thefts of parts from inside or outside vehicles, domestic and commercial robbery, domestic and commercial burglary, thefts from person, thefts from work, thefts from school, thefts from shops, other thefts, and criminal damage of cars or other objects.

[47]Alternatively, one could use the variable "ever committed a property crime", which would decrease the zeroes of the dependent variable by much. However, this is not an appropriate variable to use because of two reasons. Firstly, the possibility of response error would be much higher if we considered the whole lifespan of an individual, since this is a retrospective survey. Secondly, some crimes that immigrants report are committed long before the decision of an immigrant to immigrate in the host country.

[48]Violent crime consists of: assaults with and without injuries, and commercial and personal robberies.

as burglary, robbery, criminal damage, vehicle thefts, other thefts, etc). The large number of zeros, however, does not allow to use these crime types separately. In addition, since property crime involves some trivial crimes, a variable "serious property crime" could be used. However, the very low variation in the serious property crime variable (98% of zeroes) was not enough to identify all coefficients of interest.

Although many respondents' characteristics are available, only controls for basic demographic characteristics, such as age, gender, regions and ethnic background are considered in the empirical investigation. This followed strategy is attributed mainly to two reasons. Firstly, this study is aiming to identify whether immigrants' criminal behavior (mainly for property crime) would be different from natives' one if immigrants and natives shared the same basic demographic characteristics. Of course, it would be interesting to explore the behavior of the impact of immigration on criminal behavior once other controls such as education, working status, parental characteristics, marital status, risk factors, etc, are included. However, most of these variables are derived from questions that involve only people older than 17 years old, which results in reducing the sample by around 2,500 individuals. Moreover, some other variables, such as risk factors contain many missing cases which would reduce the sample size even more.[49] Thus, the reduction in the sample size is the second reason why these controls are not used. Actually, the empirical investigation showed that when the estimators that control for misreporting are used, the variation of the reduced sample does not allow identification of the parameters of interest.[50] Instead, an "open" discussion will try to identify the factors that result in potential estimated crime differentials between immigrants and natives.

Proceeding now to the independent variables, it would be proper to first discuss the main regressor, a dummy that indicates whether someone is an immigrant or a native. While it is common in empirical studies to define an immigrant as the person who is not

[49]Some of the independent variables include many missing values, ranging from 180 to 424 cases. Dropping all these missing cases would result in loosing around 1,300 extra cases. Instead, dummy variables, which take the value one if the associated variable has a missing value, could be constructed and used together with the "parent" variable. However, the use of these variables seems impossible, since in most cases the variation between the regressand and these "missing" variables is very low.

[50]Actually, to achieve convergence we had to impose many exclusion restrictions from both processes, which led to bad misspecification of the models, since some of the excluded variables actually belong to the two processes. Thus, using these models would not be enough to shed light on the question of interest, and possibly it would result in misleading conclusions.

born in the host country, country of birth is not available in this survey. The question used to construct immigration status is the following: "Can I just check how long have you lived in the United Kingdom?" Respondents that replied with "All my life" are considered as natives. Otherwise, they are classified as immigrants.[51] A limitation of this construction is that there can be some natives who had left UK but returned back after a certain period of time. These people may have categorized themselves as living in the UK less than their whole life, therefore as "immigrants", although they must be considered as natives, particularly if the period of staying outside the UK was very small. People born in the UK but lived most of their life in another country would exhibit very common characteristics with immigrants. Thus, it would not be very unreasonable to place them in the same group with actual immigrants. Nevertheless, I would not expect this number to be large enough, as according to the core sample, the weighted percentage of people who did not live in the UK their whole life is 9.2%, which is quite close to the percentage of immigrants in the UK from other sources in 2003.[52] Although in the initial core sample only 729 immigrants appear, I have increased their number to around 2,000, by exploiting the young boost and most importantly the ethnic minorities boost. This has been done mostly to increase precision of "immigrant" estimates. To restore representativeness, as described shortly before, a combined weighting variable of the weights of the three distinct data sets is used.

In the remainder of this section a description of the other covariates that are used in the regression analysis is presented. These variables involve very basic characteristics of the population, such as age, gender, and region of residence. Descriptive statistics of the independent and the dependent variables can be found in Table 6.2.

It is well known in criminologists' research that age is closely linked to criminal behavior (see for example, Farrington, 1986). Most evidence suggests that crime peaks in the teenage years and then falls steadily. Therefore, higher powers for the "age" variable should be also used. Gender is another very significant determinant of crime as men's crime rates are

[51] Actually, respondents had to choice among the following alternatives: 1) Less than 12 months, 2) More than 12 months but less than 2 years, 3) More than 2 years but less than 5 years, 4) more than 5 years but less than 10 years, 5) 10 years or more, and 6) All of his or her life.

[52] For instance, according to the OECD estimates, the proportion of foreign born population in the total population in the UK was 8.8% in 2003 and 9.3% in 2004.

universally much higher than women's (see for example, Steffensmeier and Allan, 1996). Thus, a dummy distinguishing men from women is used. Finally, controlling for the region where the respondent lives seems to be quite important to capture regional unobserved characteristics associated with crime (see for example, Glaeser and Sacerdote, 1999), such as high poverty rates, or high unemployment rates. Moreover, different areas may be associated with higher returns to illegal acts, or, lower probabilities of arrest. Using the standard regional variable, ten regional dummies have been constructed. Nonetheless, the very demanding econometric models require the grouping of regions into four groups. These are North (North, York, North West), Midlands (East, West, and Wales), and South (East Anglia, South East and South West). London will be the baseline group.

In a second specification also ethnicity (or race) is added. Criminologists have devoted much research on how different ethnic groups relate to different criminal behavior (see for example, Torny, 1997). Most of them find that individuals from ethnic minorities are disproportionately represented in official crime records.[53] However, using self-reports the opposite is found (see, for example, Sharp and Budd, 2005). The reasoning behind any possible link between different ethnic groups and crime, shares many arguments similar to the link between immigration and crime. Since a higher fraction of immigrants come from ethnic minorities relative to the native population, we would expect that inserting dummies for ethnic groups would have a strong effect on the impact of immigration status on crime.[54] Five dummies for ethnic groups are constructed. White or not, Black or not, Asian or not, Mixed or not, and Other Ethnicity or not. As the proportion of people of other ethnicity is too low, Other Ethnicity is grouped together with Asians. This does not seem inappropriate as around 50% of people from other ethnic groups are Chinese.

Finally, I would like to devote a few lines on two "special" variables which will be also used in the next section. As described in Section 5 identification of count data models requires at least one exclusion restriction on crime process, or a sign restriction on the reporting process. Although, as we will see at the next section, there is no available *a priori* information on any sign of the parameters in the reporting process, we can use

[53]It must be stressed that ethnic minorities are most of the times defined as non-white individuals.

[54]In the core sample, the fraction of immigrants to the total population of ethnic minorities is about 61%.

some information in the data set to impose an exclusion restriction on the crime process by constructing a variable assumed to belong to the reporting process only. Note that, as was also described in Section 5, the binary choice models are identified even without any exclusion restriction. However, this extra information will be also used in binary choice models in order for the analysis to be consistent across all models, and because it can facilitate the estimation procedure.[55] In the next section two available options are described.

Firstly, respondents have been asked whether they replied to the questions concerning crime truthfully. Thus, a dummy variable of truthfulness can be generated. As will be explained also later, it is reasonable to assume that this variable belongs only to the reporting process, as can be considered as a characteristic that shapes the reporting behavior. Therefore, since this variable appears only in the reporting process, we technically have an exclusion restriction on the crime process. This will help to distinguish the probability to report a crime from the probability to commit a crime, even though reliability of responses on this question is also doubtful.[56]

Secondly, although interviewers tried to provide a private environment while conducting the interviews, in 32% of the cases (3,768 cases) there was someone else present during the interview. Even though crime questions in the OCJS are self-completed in a computer (as opposed to face-to-face interviews), and although it was stressed by the interviewers that nobody should disturb the interviewee during the self-completion part, it is still possible that presence of someone else could affect the reporting behavior of the respondents. Therefore, a dummy is constructed that takes the value 1 if someone else was present. Since there are 409 missing cases, I have also constructed a dummy that takes the value one for these missing cases. Moreover, in the cases where someone else was actually present during the self-completion part, there is the extra information whether this other person actually looked at the screen (15% of the 3,768 cases). Thus, a dummy variable is generated to capture the fact that the reporting behavior might have been more affected in the cases

[55] However, as will be made clear in the next section the exclusion restriction does not drive the binary models results.

[56] It is noteworthy that 93% of the respondents replied that they truthfully answered all questions concerning crime.

where someone else looked at the answers. More discussion on these variables will follow in subsection 7.2.b.

7 Results

As discussed in Section 5, the main results follow from the binary choice variable, whether or not someone has committed a crime last year, whereas the extra information from counts is used for robustness checks. This Section is organized as follows: First, the results of conventional models for binary and count data are presented. Then, the main findings of the binary models that allow for misclassification probabilities follow in subsection 7.2. A series of robustness check results follow in subsection 7.3. Finally, in subsection 7.4 models with interaction terms of immigration status with ethnic status and regional dummies is discussed.[57]

7.1 Preliminary Results

To acquire a first idea of the impact of being an immigrant on crime, conventional Probit results of regressing the crime variables on immigration status only are presented in Table 7.1.a.[58] Even though this work focuses on property crime, in this table also results for all the different categories of crime are presented in order for the reader to obtain a more general idea. Assuming that there is no misreporting, this table clearly shows that being an immigrant does decrease the probability of committing crimes for all categories. Table 7.2.b portrays the results for the count form of these variables, using a Negative Binomial 2 model (NB2). Not only are immigrants less likely to commit crimes but they also commit fewer crimes. This is evident in these results, since taking into account the number of

[57]Throughout the empirical research the appropriate sampling weights to restore representativeness of the sample are used. This is because there are different sampling rates for young people (young Boost) and ethnic minorities (ethnic Boost). It must be stressed that if the model is correctly specified and the stratification is based on the independent variables (which is the case here), unweighted regression analysis will still result in consistent estimators (see, for example, Wooldridge, 2003). However, since weighted and unweighted regressions yield different estimates across different models, use of the weights seems more appropriate. However, as will be discussed in Section 7.3.3, the unweighted results of the covariate-dependent Probit with misclassification are very similar to the weighted ones.

[58]Although, Complementary Log-Log (CLogLog) models might be more appropriate when there are so many zeroes, as it assumes an asymmetric cumulative function, the empirical results showed that nothing was gained by using this model, in terms of better fit or any differences in the estimates. Moreover, specification tests were not of support of using an asymmetric binary model such as the CLogLog.

reported crimes, immigration coefficient becomes even more statistically significant in all but drugs related offences. As mentioned in the previous section, there are 54 extra missing cases in the count form of property crime variable. However, this is not much of a concern since the Probit results hardly change even when these missing cases are dropped.

As these models are nonlinear, interpretation requires the calculation of marginal effects. Focusing on property crime, we find that being an immigrant reduces the probability of committing a property offence in last year by 1.81 percentage points, a relative change of 46.78%.[59] Regarding the NegBin2 model, this preliminary estimation says that, without taking into account differences in demographic factors, an immigrant yearly commits 0.160 crimes on average whereas a native commits 0.366 crimes, a difference of 0.206 crimes (a percentage decrease of 77.67%).

Nonetheless, it would be more interesting to see whether immigrants exhibit different criminal behaviour than otherwise comparable, in terms of basic demographic characteristics, native-born individuals. Thus, in Tables 7.2.a-b we look at this difference, once we have controlled for fundamental demographical features such as gender, age and region of residence in specification 1, and also for ethnicity in specifications 2 and 3. Although the sign maintains, the statistical significance fades away in both binary and count models.[60] Specifications 2 and 3 show that the coefficient on immigration dummy becomes even smaller and less significant once we control for ethnicity. This is expected, since immigrants are relatively more nonwhite than natives (see, Table 6.2) and, as the results show, white individuals actually report that they are more prone to crime. However, in the NB2 model the immigration status coefficient does not lose so much of its magnitude. Finally, from specification 3 in both Table 7.2.a and Table 7.2.b we can see that Asians & Others, and in a lesser degree Blacks, are less prone to crime than otherwise comparable Whites.[61]

[59] The probability of a native committing a crime is 0.0568, whereas for an immigrant it is 0.0387

[60] Although insignificant, it would be interesting to evaluate the magnitude of these coefficients once we have constructed a "representative" individual. Thus, what would be the difference in the probability of committing a property crime between a native and an immigrant, who are both 25 years old, males, and live in London? According to this model this is -1.95 percentage points but statistically insignificant, a relative effect of 26.37%. The NegBin2 model says that being an immigrant, holding all other characteristics constant, reduces the expected number of crime by 0.12, which is statistically insignificant as well.

[61] Note that the variable 'age' and its quadratic do not provide a very good fit. To obtain the pattern of the impact of age on property crime we need to include up to the fourth power of age. Following this, immigration dummy's effect becomes slightly more significant. Here, we present only the quadratic term in order to be in line with the models that control for misreporting.

The above results are just a first indicator of the true crime picture as we can hardly draw inference for the actual crime relying on the reporting crime, unless there is no under-reporting. According to these results we are only confident to say that for some reasons immigrants report less crime, but this difference fades away once we control for some basic characteristic. However, some would argue that this difference exists because immigrants may under-report criminal activities by more than natives. If this is the case, immigrants may still commit more crimes than natives but at the same time under-report by more, resulting in this negative coefficient. This can also be argued for the case of white individuals if the groups of ethnic minorities are reluctant to report truthfully. Nevertheless, more appropriate (parametric) models that control for this possibility, resulting in coefficients corresponding to the actual crime, are presented in this Section.

7.2 Probit Model that allows for Misclassification

7.2.1 Constant Misclassification

In this subsection the results from the model of constant misclassification are presented (MisProbit). As discussed in Section 5.1,this is a parametric model that takes into account both under-reporting (misclassification of one as zero) and over-reporting (misclassification of zero as one). It is noteworthy, however, that this model assumes constant misclassification. Thus, it captures the actual probability of committing a crime, but it will be mispecified if probabilities of misclassification do depend on covariates. Although it may be sensible for over-reporting to be considered as constant, since we can assume that people may over-report randomly, it cannot be the case for under-reporting. It is highly possible that the same characteristics affect both the probability of committing a crime and the probability of reporting it. The assumption of constant misclassification will be relaxed in the next subsection.

Table 7.3 presents the results of MisProbit in three specifications, as in the previous subsection. There are a few important findings that deserve some discussion. To start with, we notice that in the 1^{st} specification, being an immigrant still decreases the probability of committing a crime. However, this coefficient is even less significant than the conventional

Probit. We can also see that this coefficient turns positive once we control for ethnicity but it is always very statistically insignificant.

It is also very important to stress the magnitude of misclassification of one as zero, which is around 81% and very statistically significant. This seems very large, as this estimate indicates that 81% of people have committed crimes but have reported none of them. However, as emphasized in the previous section, we must be cautious with the interpretation of this probability as this model cannot distinguish the probability of under-reporting from the probability of never committing a crime and therefore never reporting one (zero-inflation). Thus, a fraction of this probability consists of genuine non criminals and the rest of it can be considered as probability of total under-reporting. Nevertheless, nothing can be said about the importance of each, unless we model it somehow into the likelihood function.[62] In any way, the interpretation of the coefficients does not change, which still capture true crime given that misclassification is constant. Concerning the probability of misclassification of zero as one, it has a clear-cut interpretation as probability of over-reporting, since interpretation as one-inflation seems unreasonable.[63] Moreover, the estimated value of this probability of 0.012 is also expected, as we would not expect that people would report crimes that they did not commit. However, we cannot ignore it since it is very statistically significant.

Moreover, as was also discussed in Section 5.1, although the sample size is fairly large, the estimates are very imprecise perhaps because of the noisy nature of crime self-reports.[64] The only coefficients that preserve some of their significance are the coefficients on gender and ethnicity. We also notice that although less imprecise, the coefficients are quite larger in size. Furthermore, the average conditional probability of committing a crime calculated as $\Pr(\widehat{y_i = 1}) = \sum_{i=1}^{n} \Phi(x_i'\hat{\beta})/n$, is now around 29%. This is much higher than the predicted

[62] Such a model together with results is presented in Appendix A. Although it seems that identification of this model requires the same conditions as the simple MisProbit, it does not behave very well in estimation terms and consequently its results are questionable. This may be a consequence of the very noisy (crime self-reported) data used in this paper. Future research using less noisy data and larger/richer samples could reveal more interesting things about the behavior of this model. Also, further theoretical investigation of this model could reveal interesting outcomes.

[63] In this context, one-inflation would mean that: given the set of covariates x_i some people always commit and always report that they have committed crimes independent of these x_i.

[64] Note also, that if misreporting exists, the standard Probit model overestimates the asymptotic t-statistics (Hausman at al, 1998).

average probability of the simple Probit model, which is calculated to be 6%.

It can be also noticed that the maximum of the MisProbit model corresponds to a log likelihood value that is only slightly larger than the log likelihood of the conventional Probit, even though two extra statistically significant parameters are added into the model. This might be due to the fact that the estimated coefficients of the MisProbit model are less precise in comparison to Probit.

7.2.2 Allowing Misclassification of 1 as 0 to depend on Regressors

As opposed to the previous subsection, this subsection presents the results of a MisProbit model in which misclassification of true ones as zeroes (under-reporting) is allowed to depend on regressors, whereas misclassification of true zeroes as ones (over-reporting) is assumed to be constant.[65] Since the same individual is responsible for both actions of committing and reporting a crime, logically both processes are functions of the same variables.[66] However, as mentioned in Section 5, exclusion restrictions would assist the estimation procedure, even though this model is identified even without any exclusions (see, Hausman at al, 1998). To be consistent with the NB2-Logit model which is used as a robustness check (see, subsection 7.3.a), an exclusion restriction on the crime process is used, as this is crucial for the identification of the NB2-Logit.[67] The exclusion must be "strong", in the sense that the variable which is excluded from the crime process must have a significant effect on the reporting process. Otherwise, inserting variables that have no effect on the reporting part, and at the same time are correlated with the rest of the variables in this part, could result in undesirable outcomes. As described in the previous section, some information in the data set can be used to construct two variables that are

[65] Treating over-reporting as constant helps identifying all the parameters of interest. Although constant over-reporting is a sensible assumption to make, there might still be cases where this probability depends on the regressors. Nevertheless, the estimation analysis showed that identifying all coefficients of a fully specified model (a model that includes the same regressors in both probabilities of misclassification) seemed impossible. The estimation analysis also showed that it is feasible to identify all parameters of a model where there are extra exclusion restrictions from the reporting process (these results are available upon request). However, this practice would result in a mispecified model, since many variables that actually belong to the reporting process are excluded.

[66] For instance, age affects both the probability to commit a crime, as younger people commit more crimes in general, and the probability to report a crime, as younger individuals would be less willing to reveal their true criminal behavior. Similar arguments hold for the other independent variables.

[67] Note that exclusion restrictions on the reporting process does not solve the identification problem (see, Papadopoulos and Santos Silva, 2008)

assumed to affect only the reporting process.

A first choice would be to use the information whether someone else was present during the self-completion part. There is evidence, at least for face-to-face interviews (see, for example, Aquilino, 1993) that someone else's presence during responding to sensitive questions affects the reporting behavior. However, since the questions about crime were (computer-based) self-completed, which is a much more private environment, the effect of this dummy could be much smaller than in face-to-face interviews. Indeed, as the results show, this turns to be a "weak" restriction.

In another direction, the variable "truthfulness" can be exploited, a dummy that takes the value one if people said that responded in all questions concerning crime truthfully. This variable is used only in the reporting process, as it makes sense to assume that whether or not someone has truthfully reported his actual criminal activity at the time of the survey, could not affect the action of committing a crime before the survey took place. If any empirical relationship exists, this would be because "truthfulness" is correlated with unobserved characteristics correlated with crime, or because there is a reverse causality of committed crimes on "truthfulness".[68] The problem is that it is also not correct to assume that "truthfulness" actually affects the reporting behavior, unless the reported "truthfulness" coincides with the actual behavioral characteristic of how truthful someone is. However, what we assume here is that being "truthful" while answering questions about crime is a feature that "shapes" some behavioral attributes, which in turn affects the reporting behavior. In any way, as will be discussed also later, the results show that "truthfulness" actually has a very significant effect on the reporting process but no effect on the crime process, once "truthfulness" is included in both processes.[69]

According to all the above, the main results of this section are based on the specification where "truthfulness" affects only the reporting process. In the robustness check section it

[68]For example, the probability to answer "I was truthful" would be higher for people who commit more crimes but report fewer, if this was a way to hide misreporting. Or, it might be that, it is less possible for people who commit no crimes to say that they are not truthful, as there is no reason for them to lie. In both cases we would expect a negative relationship between reported crime and "truthfulness". In fact, a weighted Probit regression of "trufulness" on number of reported property crimes, showed that this is actually the case.

[69]This is true for both binary choice, and count data models. If "truthfulness" is included only in the crime process, it has a small but statistically significant effect.

will be shown that this exclusion restriction does not drive the results. Some results of using "other's presence" as an exclusion restriction are also presented in the robustness check section. There, we will also see that the inclusion of this dummy has some undesirable effects on the estimation of the model. As we discussed in the previous section a dummy "someone looked at the screen during the self completion part" can be also constructed. The results also show that this information is unrelated to the probability of not reporting committed crimes.[70]

The results of this model are depicted in Table 7.4. The estimated coefficients of the crime process are reported in the upper part of this table, whereas the coefficients of the "reporting" process are presented in the lower part. Before discussing the effect of the immigration dummy, I would like to mention some main features of the findings of this model. First of all, the log likelihood corresponding to the global maximum is considerably improved comparing to the previous model. Since the covariate-dependent MisProbit model nests the MisProbit with constant misclassification (when all coefficients but the constant of the "reporting process" are zero), we can construct a likelihood ratio test to test whether misclassifying one as zero is constant. Since the likelihood ratio statistics is around 49 for all three specifications, there is strong evidence against the null hypothesis that misclassification of one as zero is constant.[71] Moreover, the predicted average probability of committing a property crime during the last period, calculated as $\sum_{i=1}^{n} \Phi(x_i'\hat{\beta})/n$, is around 29%, which is in line with the result of the previous subsection.[72] However, the average probability of misclassifying an one as zero, calculated now as $\sum_{i=1}^{n} \Phi(z_i'\hat{\gamma})/n$, is 10 percentage points lower than the previous model. Finally, notice that the coefficient of over-reporting is again around 1.3% and statistically significant at 1%.

Concerning the main objective of this research work, this table shows that even after controlling for any potential difference in the reporting behavior of immigrants, the coefficient

[70]This dummy has no effect either using it as an interaction term (so that conditional on the presence of someone else there is no effect of some of them looking at the screen), or using it alone without controlling for the cases where someone was present but did not look at the screen.

[71]A wald test that all coefficients of the "reporting" process but the constant are zero gives similar results.

[72]Note that, although the estimates of the BCS are not directly comparable to the ones from the OCJS we have estimated that the probability of suffering a crime in 2008 was around 0.2. Notice finally, that in BCS commercial crimes and crimes against children are not included, and that property crime was higher in 2003.

on immigration status is still negative and fairly larger than before. In the 1^{st} specification this coefficient doubles in size, comparing to the previous model, and becomes statistically significant at 10% significance level. However, after controlling for ethnicity, although still negative and fairly large, it becomes insignificant. Finally, from all three specifications we can say that being the "representative" individual and immigrant, reduces the probability of committing a property crime by around 6 percentage points, before controlling for ethnicity, and around 4 percentage points, after controlling for ethnicity.[73]

The reason why immigration status coefficient becomes larger in magnitude can be attributed to the fact that native-born individuals in fact under-report by more than immigrants, and not the opposite. However, as mentioned before, the coefficients on the reporting behavior can also take a zero-inflation interpretation. Therefore, the negative coefficient might also mean that a smaller proportion of immigrants belong to the group of genuine non-criminals. These two interpretations contradict each other but we cannot say which effect is larger.[74] If the "reporting" process was measuring only under-reporting, it would be easier to analyze what would be the direction of the change in crime process coefficients once we control for the corresponding characteristics in the "reporting" process. Thus, if immigrants under-report less, this would result in the coefficient of crime process to become more negative. However, the effects of changing the portion of zero-inflation on the crime process coefficients are not clear. According to this interpretation, a positive coefficient for immigration status would just mean that fewer immigrants relative to natives participate in the binary choice decision of committing crimes or not. Thus, it does not give information about how the remaining individuals, who may or may not commit crimes, behave. In any case, we must say that the total effect of controlling for under-reporting on the coefficients in the crime process is not easy to predict, since this will depend on all inter-correlations of the coefficients across the two processes and all coefficients within a process. However, notice that the coefficient of immigration status on the reporting process

[73] The predicted probabilities to commit a crime for the "representative" individual are 4.27%, and 9.87%, for an immigrant and a native, respectively. This corresponds to the marginal effect of 5.6 percentage points, and relative effect of 131.14%. After controlling for ethnicity the above figures become $\approx$ 6%, 10%, 4 percentage points, 66.7%, respectively.

[74] If we accept the zero-inflation interpretation the negative coefficient means that the number of genuine non criminals is relatively higher for natives.

is statistically insignificant in all specifications.

It would be also interesting to briefly discuss the effects of the other explanatory variables. To begin with, it is notable that being a white individual still increases the probability of committing a crime. However, this effect is only significant at 10% significance level. From the 3^{rd} specification, we notice that black individuals' coefficient is negative and significant at 10%, in contrast to the previous model where it was very insignificant. On the other hand, the coefficient on "Asian & Others" dummy is still negative but now insignificant.

Concerning gender, the sign on "male" dummy is still the expected one since males commit more crimes than females. Males' coefficient is still significant at 1% significance level, even though the sign in the reporting process is negative and significant at 10%. This negative sign indicates that females are more reluctant to report their criminal activities truthfully, perhaps because of "embarrassment" effects.[75] Also, this may also indicate that it is more likely for a female to belong to the genuine non criminal group of people, which is also reasonable.

The results of the regional dummies are also interesting. First of all, including these dummies in the reporting process, we control for area-specific unobserved characteristics that may affect the decision to misclassify. We see that people who do not live in London commit more crime. Comparing to the previous models, the magnitude and significance of these three coefficients increase. Nevertheless, only people living in North England seem to commit significantly more crime (but just in 10%). This change may be attributed to the significant effect that this dummy has on the reporting process.

Furthermore, we can see that both age and age squared have a significant effect in both processes. Concerning the crime process, the negative sign on age and the positive sign on age squared variables indicate that crime falls with age but in a decreasing rate. The very small value of "age^2" shows that the crime reaches a minimum at advanced ages, and then increases again. We must notice that these results do not coincide with the theoretical views of the effect of age on crime, where crime increases with age during the adolescence years and then steadily falls. Although for conventional Probit this shape is captured once

[75] It is relatively less acceptable by the society if a womam commits a crime.

age, age^2, age^3 and age^4 are included, for covariate-depended MisProbit, including higher powers of age does not affect the results.[76] The coefficients of age and age^2 variables in the reporting process are also negative and positive respectively. This also predicts that people under-report by less as they are getting older at a decreasing rate (or that they are increasingly likely to switch to the category of potential criminals).

Special attention must be finally paid to the effect of the "truthfulness" dummy. As mentioned before, this dummy belongs only to the reporting process.[77] Table 7.4 shows that this dummy has the largest coefficient in the reporting part. It is also statistically significant at 1% significance level. This coefficient can be interpreted in two contradictory ways. Respondents who replied that they answer all crime questions truthfully are honest people who, regardless the other observed characteristics, never commit and never report crimes. On the other hand, this coefficient may also indicate that "truthful" respondents are people that under-report more than "non truthful" respondents, so that they commit more crime than what they actually report.

7.3 Robustness Checks

7.3.1 Count Data Models

In this subsection we examine whether the results of modified count data models that take into account under-reporting coincide with the main findings of the MisProbit. Of course these two models are not directly comparable, since the count data models use the extra information of the number of property crimes. Nevertheless, similar estimates between the count and the binary models used in the present study, mostly with respect to the reporting process, would strengthen the reliability of the results of the main analysis. As described in the previous section, the unconditional variance of the dependent variable is much larger than the unconditional mean. This is a first rough but strong indicator

[76]Regressions that include up to the 5^{th} power of Age have been performed. Including more powers, results in no convergence of the optimization procedure.

[77]The estimation results, when "truthfulness" is also included in both processes of the covariate-dependent MisProbit model show that this dummy has no effect in the crime process, and do not change the results. Thus, these results also support the theoretical assumption that "truthfulness" should not be included in the crime process. However, even if there was an effect, this would be misleading since a significant effect would just capture unobserved characteristics that are both correlated with crime and how truthful someone is.

against the equi-dispersion assumption of the Poisson distribution that the mean equals the variance. Therefore the NB2 distribution, that allows for over-dispersion by using an extra parameter that accounts for extra unobserved heterogeneity, may provide a better fit to the data. Table 7.5 portrays the results of three NB2 models. The 1^{st} column reproduces the results of the simple NB2 model for the sake of comparisons. The 2^{nd} model, which is the NB2-Logit, controls for under-reporting, while the 3^{rd} one also incorporates Zero-Inflation (the last model is presented in the Appendix B).

Regarding the NegBin2-Logit model, Papadopoulos and Santos Silva (2008) showed that unless exclusion restrictions are imposed on the count part, there are two linearly dependent sets of parameters that correspond to the same maximum likelihood value. Consequently, the model is unidentified since it cannot be said which set of estimated parameters is the "correct" one. Although sign restrictions in the reporting part is a possible solution, here, there is not any *a priori* information to suggest the sign of any of the parameters of the reporting process. Therefore, in line with the binary choice models, to identify all the parameters of the model, the "truthfulness" dummy is used in the reporting process only. In subsection 7.3.d we will show the consequences of excluding the dummy "other present" from the crime process. However, we must be very cautious since, although the model is globally identified, there can still exist more than one maxima. Therefore, a thorough analysis must be performed to find all possible maxima.[78] Also notice that this model does not control for over-reporting, differing in this aspect from binary choice models. However, the covariate-dependent MisProbit gives very similar results even when the probability of over-reporting is assumed to be zero.[79] It is also important to stress that the structure of this model provides more information about the data generating process than the binary choice model. The binary models only provide information about reporting or not crimes, regardless of how many crimes someone has committed. This model on the other hand, provides estimates for the probability of any given committed crime to be reported.

The regression analysis showed that the global maximum of the NegBin2-Logit is the

[78] Some tips on several possible ways to find the best maximum are described in Papadopoulos and Santos Silva (2008). In the present analysis, the regression analysis showed that several local maxima exist.

[79] These results are available upon request. This similarity of the coefficients across the two models was expected, since the probability of over-reporting is too small to affect the parameters of the other processes.

one depicted in Table 7.5.[80] The upper part of this table presents the coefficients of the crime, or differently, count process. The lower part presents the coefficients of the reporting process, which is the probability of reporting a committed crime. First of all, the very large value of α must be noticed, which is statistical significant in any significance level.[81]. Therefore, there is evidence that the data are over-dispersed even after conditioning on the set of regressors.

Regarding the immigration status coefficient in the crime process, we can see that even after controlling for under-reporting, it is negative and even larger in value than the conventional NB2 model. This is the consequence of immigrants' coefficient in the reporting process being positive. That is, being an immigrant increases the probability of reporting a given crime and therefore decreases the conditional expectation of crime by more than the conventional NB2 model. This finding is consistent with the binary models. In addition, in line with the binary models, the coefficient of immigrant dummy is insignificant in both processes. I would like to mention that in NB2-Logit, contrary to the covariate-dependent MisProbit, the immigration status coefficient does not depend on whether or not we control for ethnicity. This is the reason why only the model that controls for ethnicity is presented in this subsection. Finally, the marginal effect of our "representative" individual says that being an immigrant decreases the expected number of crimes by around 0.19 crimes (this difference was 0.11 for the simple NB2 model).[82]

As far as the rest of the coefficients in the crime process is concerned, their direction is in accordance with the coefficients of the MisProbit model. However, we must stress that the NB2-Logit models the conditional mean of crime events, whereas the MisProbit models the conditional probability of committing a crime. Therefore, it is always possible that few

---

[80]The estimation analysis showed that another maximum exists with log likelihood value of 2,315.80. This maximum corresponds to coefficients very different from the global maximum. These results are available from the author upon request. As shown in Papadopoulos and Santos Silva (2008), there is a relationship between the coefficients of the two models. A brief description of this relationship follows in subsection 7.3.4 Although the log likelihood value of this local maximum is close to the log likelihood value of the global maximum, the difference is sufficient to permit identification of the "correct" maximum depicted in Table 7.5. This is because the excluded variable "truthfulness" has a strong significant effect on the reporting process. In subsection 7.3.4 we show what are the consequences of a "weak" exclusion restriction.

[81]The log likelihood of the corresponding global Poisson-Logit maximum is -9,132.31

[82]The expected number of committed crimes by the "representative" individual are, 0.1771, and 0.3628, for an immigrant and a native, respectively. This corresponds to the marginal effect of 0.19 crimes, and the relative effect of 104.9%.

differences exist across binary and count data models, even if both models are correctly specified.[83]

Regarding the reporting process, this model predicts that the average conditional probability of reporting a committed crime is 43%. However, we need to distinguish this probability from the probability of under-reporting in the binary models. The MisProbit estimates the probability of reporting zero crimes given that some crimes have been committed (or the probability of never committing and consequently never reporting crimes), whereas the NB2-Logit model estimates the probability of someone reporting a committed crime, regardless of the number of crimes he has committed. Therefore, this model is in a sense more structural, since it also captures cases where people under-report but still report some of their crimes. On the other hand, the binary model ignores this kind of under-reporting since it can capture only reporting zero crimes. We can see that the sign of most of the coefficients of the Logit part are opposite to the ones of the reporting part of the MisProbit model. This is expected, since here we measure probability of reporting a committed crime, whereas in MisProbit we were measuring probability of not reporting crimes.[84]

In the 3^{rd} column, the results of the ZI-NB2-Logit model are presented (which model is presented in Appendix B). According to this model, some people never report crimes just because they never commit crimes, or because they not report any of the committed crimes. Therefore, the zero-inflation probability, in line with the MisProbit, gives the proportion of people that totally under-report or the proportion of genuine non-criminals. The rest of the individuals may or may not commit crimes, but their responses are still subject to under-reporting. In this sense, we assume that not everyone is a potential criminal, so that not everyone participates in the choice whether or not to commit crimes. In other words, the Logit part of this model measures the probability of reporting a committed crime once

[83]For example, here the coefficient of "white" dummy is insignificant but it was significant at 10% significance level in the MisProbit model. This might mean that although white people are more likely to commit a property crime, they do not commit more crimes than nonwhite individuals, so that taking into account the extra information of the counts reduces the white-nonwhite crime differential.

[84]However, I would like to repeat that in MisProbit we could not separate zero-inflation from under-reporting, and therefore exact comparison between the coefficients of the two models would not be appropriate. Anyhow, the only striking difference is that the NB2-Logit model predicts that being a male decreases the probability of reporting a given crime, while in the MisProbit being a male decrease the probability of reporting no crimes, either because of misclassification or because of never committing and consequently never reporting crimes.

we partial out people that always under-report with zero, or people that never commit and consequently never report crimes.

Since identification of this model requires the same conditions as in NB2-Logit, the exclusion restriction of "truthfulness" from the crime process is also followed. To model the zero-inflation process I make use of the same variables I use to model the crime process. The coefficients of the ZI-NB2-Logit model that correspond to the global maximum are presented in the last column of Table 7.4.[85] An interesting finding of this model is that the conditional predicted probability of zero-inflation is around 62%. This is very close to what the covariate-dependent MisProbit predicted. There, the same probability was calculated to be around 71%.

This model also predicts that, even after controlling for zero-inflation, the probability of reporting a committed crime is 37.6%, which is even lower than in NB2-Logit model. According to this model, once we control for zero-inflation and under-reporting, immigrants' coefficient decreases even more. However, it is still statistically insignificant as the precision of the estimate decreases. We can also notice that in ZI-NegBin2-Logit, males' coefficient becomes insignificant. More interestingly, the coefficient of "white" dummy turns negative. This is perhaps because of the negative and statistically significant coefficient of this variable on the zero-inflation process, which may mean that the probability of white individuals to always under-report with zero is smaller than non-white persons. Finally, we notice that people who live in South and in North commit more crimes than people who live in London (significant at 5% and 1% respectively).

Since the "zero-inflation" process of the ZI-NB2-Logit has the same interpretation as the "misclassification of one as zero" process of the covariate-dependent MisProbit, it would be interesting to compare the corresponding coefficients of the two models. We notice that apart from the coefficients on "age" variables, all other coefficients follow the same direction. Furthermore, it can be observed that there are a few differences in the statistical significance of some coefficients. Finally, regarding the coefficients of the "reporting a committed crime" process, they are relatively similar in both models.

[85]As was the case for NB2-Logit, the estimation analysis shows that, again, another maximum exists with log likelihood value of 2,261.28. These results are available from the author upon request.

7.3.2 Violent Crime

In this subsection we briefly investigate what is the relationship between immigration and violent crime. Although violent crime is more impulsive, some of the reasons used in Section 2 can be also applied here to hypothesize a link between being an immigrant and committing violent crimes. For example, according to Merton's (1938) "strain theory" immigrants may become violent due to accumulation pressure because of discrimination, or racist behavior against them by native population. On the other hand, a credible behavior associated with "no crime" would be a good path of integration in the host country. Furthermore, risk attitudes and deterrent effects, which are very important in explaining both the decision to commit property and violent crimes, might be different between immigrants and natives. Other reasons that act in the one or the other direction can be thought of. Therefore, as was the case with property crime, empirical investigation can offer more insights on this link.

The results, presented in Table 7.6, are obtained using the MisProbit model. The same specification as the 2^{nd} specification of Table 7.4 is used. In the 1^{st} column the results of property crime are reproduced, whereas in the 2^{nd} column the violent crime results are depicted. It is striking how close the results between the two models are, since apart from the coefficients of the regional dummies, all other coefficients are very similar.[86] We can see that the same basic demographical characteristics are good predictors of both violent crime as well as property crime. Also, we see that the probabilities of committing a violent crime but not reporting it are lower than for property crime. Finally, separating ethnicity in three groups as before, I find that now Asians & Others is the least crime-prone group.

Concerning the effect of the immigration dummy, it is again negative but slightly less significant. Hence, immigrants are slightly more law-abiding than natives for both crime types. This might be because immigrants are more risk averse, or because they are more responsive to deterrent effects.[87]

[86]The tetrachoric correlation coefficient (see, Edwards and Edwards, 1984) is 0.5760, so that it is not the case that the results are too close just because the same people who committed property crimes also committed violent crimes. In addition, notice that although both crimes include robberies, this type of crime accounts only for a very small proportion of the total number of property or violent crimes (1.2% for property crime and 1.1% for violence.

[87]Finally, use of the interaction terms, showed that immigrants do not interact with the observed covariates in the same way as in property crime (see, subsection 7.4). In particular the interaction terms of

7.3.3 Weighted *versus* Unweighted Regressions of Property Crime

As noted at the beginning of Section 7, throughout the empirical research the appropriate weights to restore representativeness of the sample are used. However, if the model is correctly specified, both weighted and unweighted estimators are consistent, but unweighted ones are also more efficient (see, Wooldridge, 2003). Thus, if the estimated parameters of the unweighted model are very close to the parameters of the model that uses weights, there is some support of correct specification of the model.

According to my results, the weighted estimates are very different from the unweighted estimates in the constant-misclassification MisProbit model.[88] It is noteworthy, however, that the coefficient values of the weighted estimates of the covariate-dependent MisProbit are very close to the unweighted ones (see, last column of Table 7.6). Moreover, it is evident that the coefficients of the unweighted regression are more precisely estimated. The only notable difference is that in the unweighted estimation the effect of immigration is higher in terms of magnitude and statistically significant at 5%. This might be the case because the unweighted estimator is more efficient, so that the immigration coefficient in the weighted estimation is less precisely estimated. Furthermore, as we have included an ethnic boost data set, immigrant population is over-represented in my sample size. Thus, using weights has as a result to down-weigh the immigration sample, which may induce differences in immigration status coefficient. Note also that among all variables used in these models the immigration dummy is the variable with the most zeroes. Hence, down-weighing (over-weighing) a variable with a low variation could result in higher differences between weighted and unweighted regressions, than down-weighing (over-weighing) a variable with higher variation. For instance, we can see that although the use of the weighting variable down-weights young people (as we use a young boost), the differences of the age

immigrant with regions, immigrant with gender, and immigrants with ethnicity are not important at all. Finally, for unidentified reasons when interaction terms of age and age2 are included, convergence cannot be achieved.

[88] Although the differences in the coefficients between conventional weighted Probit and unweighted Probit are smaller than the differences between weighted constant-misclassification MisProbit and unweighted constant-misclassification MisProbit, this cannot be an argument that the conventional Probit model is a better specified model than the MisProbit. This is because Probit is designed to capture reported crime but MisProbit intends to capture actual crime. What these results might indicate, is that a Probit model is probably a correct specification for reported crime, but constant-misclassification MisProbit is an incorrect specification for actual crime.

variable coefficients are marginal. Finally, notice that the biggest differences are observed in the coefficients of the variables that are insignificant when we use weights, such as the coefficients of the regional dummies.

7.3.4 Are the Results Driven by the Exclusion Restriction of the "Truthfulness" Dummy? *Consequences of using "Other Present" instead of "Truthfulness"*

In this subsection I briefly intent to show that the main results are not driven by the use of the exclusion restriction "truthfulness". Indeed, the covariate-dependent MisProbit gives quite similar results even without any exclusion restrictions. Moreover, we briefly examine the consequences of using the dummy "other present"as an exclusion restriction, which seems to be a rather weak restriction for property crime but a "strong" restriction for violent crime. On the other hand, exclusion restrictions are necessary to identify both the NB2-Logit and the ZI-NB2-Logit models. Therefore, also the results of using "other present" as a dummy in count data models are discussed. It needs to be stressed that when the dummy "other present" is used, as was explained in Section 6, a dummy for its missing cases must be also included.

First of all, the results of the covariate-dependent MisProbit are discussed, which are given in Table 7.7.a for property crime and Table 7.7.b for violent crime. The 1^{st} specification of Table 7.7.a and Table 7.7.b reproduces the 1^{st} (for property crime) and 2^{nd} (for violent crime) specifications of Table 7.6 respectively, where the variable "truthfulness" is used. For both property and violent crime we can clearly see that the coefficients of the 2^{nd} specification, where there is no exclusion restriction, are fairly similar to the coefficients of the 1^{st} specification. Concerning the coefficient of main interest, we can see that for both property and violent crimes, the probability for an immigrant to commit a crime slightly increases, but it is still negative and similar in terms of significance. Thus, according to these results, the use of "truthfulness" does not affect the main results of the model.

In specification 3 of Table 7.7.a, we look at the consequences of excluding "other present" from the property crime process. However, it seems that this dummy has no effect on the reporting process of property crime. Thus, technically the restriction is quite weak. We can see that inclusion of "other present" actually results in much less precise estimates for most of the parameters.[89] Thus, not only has this dummy no effect on the probability to under-report, but its interaction with the other variables in the reporting process also worsens the general behavior of the model.[90] Consequently, the effect of immigration status dummy, which is the case for the most of the variables, decreases in both magnitude and significance. However, it still retains its sign. On the other hand, it occurs that "other present" has a significant positive effect (at 1% significance level) on the reporting behavior of violent crime (it increases the probability of misclassification of one as zero). Contrary to property crime, it is noteworthy that all the estimates of this specification are very close to both the 1^{st} and 2^{nd} specifications of Table 7.7.b, in terms of both precision and magnitude. Again, the magnitude of immigrants' coefficient slightly decreases but so does the standard error, leaving significance almost unafffected. Thus, all together, there is some evidence that the results of the covariate-dependent MisProbit are robust in relation to the exclusion restriction, as long as this is a "strong" restriction.

Regarding count data models, in Table 7.7.c. I present results of including "other present" in the reporting process of NB2-Logit model to test whether the results of count data models are also robust in relation to the exclusion restriction.[91] The 1^{st} column reproduces the results of the 2^{nd} column of Table 7.4 for the sake of comparisons. As can be seen from the 2^{nd} column, "other present" has again no effect on the probability of reporting a committed property crime. However, in this case we must be very careful as another maximum very close (in terms of the log likelihood value) to the global maximum exists, which corresponds to very different parameter estimates. This is presented in the 3^{rd} column. As Papadopoulos and Santos Silva (2008) show, it appears that there is a close relationship between the parameters of the two maxima. Given that $\theta = (\beta, \gamma)$ is the set of true parameters of the model, where β corresponds to the vector of parameters of true crime and γ to the vector of the probability of reporting a committed crime, if the exclusion restriction is "weak", another maximum very close to the true one exists with

[89]Note that, if the variable for the missing cases of "other present" is not included, the precision of the estimates increases, although it is still worse than the other two specifications.

[90]If we include "other present" in the crime process as well, regardless whether we include dummies for the missing cases, we obtain much more precise estimates which are fairly close to the 2^{nd} specification.

[91]Results of ZI-NB2-Logit, which are also available from the author on request, are very similar.

parameter value $\tilde{\theta} \simeq (-\beta, \beta + \gamma)$. The stronger the exclusion, as for example the case for "truthfulness", the easier it is to distinguish the correct maximum and the higher the deviation of $\tilde{\theta}$ from $(-\beta, \beta+\gamma)$ becomes.[92] Despite the identification problem of this model, given that we accept that the correct maximum is the one in column 2, it is clear that the estimated parameters are very similar whether we use "truthfulness" or "other present" as an exclusion restriction. Once more, the coefficient on immigration status slightly reduces in magnitude but it gains in precision, resulting in a slightly higher p-value.

Note finally, that the dummy "someone else looking at the screen during the self completion part" has no effect neither in binary nor count data models for property crime, regardless whether we use it alone, or when we use two dummies, "other present but did not look" and "other present but looked". However, both dummies have an equal, statistically significant (at 5%) impact on violent crime.

Thus, according to the analysis of this subsections, the results from both binary and count data models are not driven by the exclusion restriction of "truthfulness".

7.3.5 Other Robustness Checks

In this subsection three extra robustness check are considered.

First: Notice that the dependent variable includes criminal damage (98.5% of zeroes). Even though this is also a crime against the property, it entails only psychic gains to the offenders and therefore, it is not very clear whether it is proper to include it in the property crime variable. What happens if excluded? "Immigrant" coefficients reduces a bit but still preserves its sign and some of its magnitude -0.162 (.278), but if we do not include truthfulness it increases to -0.200 (0.306). Notice though that not considering criminal damage reduces the variation of property crime from 5.47% to 4.91% and it affects some of the other coefficients as well.

Second: We are interested only in the criminal behavior of immigrants in the host country and not their criminal behavior before moving, as their countries of origin may exhibit very different characteristics associated with property crime, such as economic op-

[92] Actually, there are cases where the model identifies as global maximum the "wrong" maximum, where the parameters of the reporting process have the opposite from the expected sign. However, from the author's experience, this happens only in cases where the exclusion is very "weak".

portunities and deterrent factors. However, in my sample I also include immigrants that have been in the UK for less than 12 months. Additionally, the crime question concern the individuals' criminal behavior during the 12 months prior to the interview day. Thus the most recent immigrants may have committed crimes in their source countries and mistakenly recorded them as happened in the UK. Moreover, crimes that did not happen in the UK are not recorded. Thus, if the most recent immigrants committed crimes that were not recorded, the immigration coefficient would be downward biased. Therefore, for both reasons, I exclude this group to evaluate the effect of those immigrants who have been in the country for more than 12 months. (sample reduces to 11,541 people). Now the coeff becomes more negative, -0.302 (.287). The remaining immigrants are only very slightly older (38.69 years old without most recent immigrants, and 38.1 for all immigrants).

Third: The analysis includes people from 10 to 66, so responses of 10 year olds might be less reliable. What is the consequence of dropping very young children? Notice that the youth boost (10-25) includes 3,185 people.

Dropping 10 year olds: Immigrant -.325 (.286), sample drops by 293 people, other coeffs very similar. Prob of committing: 0.25, Prob of under: 0.66

Dropping 10-11 year olds: Immigrant -.524 (.403), sample drops by 661 people, other coeffs loose a lot of precision. Prob of committing: 0.31, Prob of under: 0.75

Dropping 10-12 year olds: Immigrant -.0.668 (.374), sample drops by 1038 people, other coeffs loose a lot of precision. Prob of committing: 0.35, Prob of under: 0.79

Dropping 10-13 year olds: Immigrant -.0.713 (.466), sample drops by 1443 people, other coeffs loose a lot of precision. Prob of committing: 0.36, Prob of under: 0.80

Dropping 10-14 year olds: NO convergence. Sample drops by 1785 people.

7.4 Decomposition of Immigrants by Ethnicity and Regions

Throughout the empirical analysis, a negative relationship between property crime and being an immigrant has been observed, other things being equal. However, we have treated immigrants as a homogeneous group of people which is not realistic. Therefore, by using interaction terms, in this subsection I decompose immigrants by ethnic groups and by region, to investigate whether different groups of immigrants are different with regard to their criminal behavior. These results are presented in Table 7.8 where the covariate-dependent MisProbit is used in all cases. Although this table presents only the estimates of the crime process, the interaction terms are also inserted in the reporting process. Thus, as in all models so far, there is only one exclusion restriction with the form of including "truthfulness" only in the reporting part. Note that all the coefficients of the interaction terms in the reporting process are statistically insignificant. All results are available from the author upon request. Finally, I would like to stress that this subsection is used to illustrate the results of the aforementioned decomposition, leaving discussion for the next section.

7.4.1 Interaction between Immigration and Ethnicity

It is a fact that immigrant population in England and Wales is very heterogeneous, as far as the ethnicity is concerned. For example, there are black immigrants coming from Caribbean or African countries, Asians from both the south and the east parts of Asia, and white population from both Europe and the "old" Commonwealth of Nations, such as Australia or Canada. Naturally, immigrants from different countries of origin have grown up with different principles, in different socioeconomic conditions, so that they differ a great deal in many aspects, both between each other and with respect to the native population. Thus, their criminal decision may differ as well. Moreover, following the same reasoning, we might also expect that foreigners who belong to an ethnic group, for instance Asians, will exhibit different behavior than natives of the same ethnic group, as the latter are better adapted in the British lifestyle.In this subsection I intend to investigate the above concepts by decomposing immigrants in four groups, which are, 'White immigrants', 'Black immigrants', 'Asian & Other immigrants', and 'Mixed immigrants'.[93]

First of all, comparing each group with the whole native population (regardless the ethnicity of natives), we find that the probability to commit a crime is considerably smaller

[93]It would be better if there was a disaggregation of immigrant population in more groups, since, for example, black immigrants from Caribbean would be different from black immigrants from Africa. The data set actually includes a derived variable that separates immigrants in 15 groups. However, the use of this variable would be impossible, because there is not enough variation between these groups and the dependent variable to identify the parameters of interest.

for black immigrants (significant at 1%).[94] Moreover, 'Asians & Other' immigrants also commit less crime than natives, but it is significant only at 20% significance level. Finally, the coefficients of the other two groups are also negative but very insignificant. From the above, it seems that there are differences in the criminal behavior among the immigrant groups. However, the only statistically significant difference is between black immigrants and white immigrants (at 10%).

Next, we investigate whether there are differences in criminal activity between each group of immigrants and their native counterparts. For this purpose, three interaction terms are used, and the results are presented in the 1^{st} specification of Table 7.8.a. From this table it seems that black immigrants commit less crime than black natives. Moreover, according to these results there is no difference in crime between the other three immigrant groups and their native counterparts. Although the interaction term "black & immigrant" is statistical insignificant, redefining the dummies by disaggregating the population in eight groups (see, Table 7.8.b) we find that black immigrants commit significantly (at 5% significance level) less crime than black natives. We also find that this is the least crime-prone group. They commit significantly less crime than all other groups but 'Asians & Others' and mixed immigrants. This is quite interesting since the involvement of black natives in criminal activities is not different than the involvement of all other groups. We can conclude that, due to some unobservable characteristics, black immigrants are less crime-prone than black, white, and mixed natives.

7.4.2 Interaction between Immigration and Region

As mentioned in the introduction, 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 their criminal activity. In this subsection interaction terms between regional dummies and immigration status are used. The results are shown in the 4^{nd} specification of Table 7.8.a.

From this table there are two things that merit some discussion. Firstly, we can notice that immigrants located in London are much less involved in criminal activities than na-

[94] These results are also not presented but they are available upon request.

tives located in London (since this is what the coefficient of the dummy "immigrant" now captures). This difference is significant at 1% level of significance. Also, we notice that immigrants who live in South England commit considerably more crime than immigrants of London. This relationship is much clearer in Table 7.8.c, where we redefine the population in eight groups according to the immigration status and the region of residence. It is clear that immigrants of London are the least crime-prone group. Apart from immigrants located in North and Midlands, they commit considerably less crime than all other groups. On the other hand, it is also interesting that immigrant population located in South is the most crime-prone category. However, they do not commit significantly more crime than the native population of Southern regions. Finally, we find that although immigrants from Midlands and North commit less crime than their native counterparts, these differences are statistically insignificant.

8 Discussion

In the previous section I presented and evaluated the results of different models that control for under-reporting. All the robustness check analysis showed that the results of covariate-dependent MisProbit are quite robust. Therefore, this discussion will be based on the results of the covariate-dependent MisProbit model. Moreover, in the following discussion we accept that the coefficients of this model actually measure true crime.

The main result of this study is that after controlling for basic demographic factors, there is a negative association between criminal behavior and immigration status. Even though the estimated difference is statistically insignificant, all the results of the sensitivity analysis section showed that it is actually quite robust. The robustness of the association indicates that this relationship probably exists, but the nature of the models and data do not allow to obtain precise estimates.

In Section 2 some channels through which there can be a positive or a negative relationship between property crime and immigration were discussed. However, it was concluded that even if these channels are "active", the final outcome is uncertain as they act in opposite directions. How can the estimated difference can be explained by the theoretical framework? A possible story is the following. It is a fact that immigrants are located in more deprived areas and confront blocked opportunities, perhaps because of human capital limitations (for example, language complexities), because employers tend to prefer natives, or due to other reasons (see, Algan, Dustmann, Glitz, and Manning, 2010). There are also, in some extent, cultural conflicts, and difficulties of adjustment. However, at the same time immigrants may be more risk averse and discount future less heavily. They might also be more responsive to the deterrent effects of potential punishment (Bucher and Piehl, 2007). In addition, they face higher probabilities of apprehension. We need to add that immigrants also face the threat of deportation, which is a punishment specific to them. Finally, coming from poorer countries, they are satisfied even with much lower economic outcomes than natives. Therefore, if we accept that some of the factors associated with more crime actually exist, we must also accept that the factors associated with lower crime work in the opposite direction over-balancing the situation. Therefore, if immigrants did not encounter the problems associated with more crime, they would be even less prone to crime than natives.

The use of interaction terms have provided some interesting insights. Although as a whole immigrants are only slightly less involved in criminal activities, it has been also found that even after controlling for demographic characteristics, immigrants in London are considerably less likely to participate in illegal activities than natives of London (but also natives of all other regions). It might be that immigrants integrate in London much easier than in other locations. Also, concentration of immigrants in specific areas might create strong social controls that discourage criminal activities. Furthermore, as immigrants are more responsive to deterrent factors, strict policing in London would discourage criminal activities of immigrants more than natives. The sure thing is that there are many unobserved cultural differences between immigrants and natives towards crime. Finally, it could be that immigrants with different criminal propensities are located in areas other than London by central agencies, such as the National Asylum Support Service. For example, asylum seekers, which is the group that according to their economic outcomes would find illegal sectors the most attractive, were located in unpopular areas outside London (see, Bell, Machin, Fasani, 2010).

But why are immigrants located in South more crime-prone than most of the other groups? This may indicate that immigrants in these areas encounter problems of adaptation into the English society, or that the socioeconomic conditions they face are less favorable than those of other regions. They may also present adverse behavior due to accumulative pressure, for example, because of discrimination, racism by natives, or cultural conflicts. Additionally, it might be the case that South England pulls the most crime-prone groups of immigrants, perhaps because there are criminal opportunities that suit them better than other groups of immigrants. It must be stressed though, that although immigrants in South are more crime-prone than immigrants in London, their involvement in crime is not statistically different from the involvement of natives in South.

Finally, we have found evidence that the group of black immigrants is significantly less involved in criminal activities than both black natives and white natives. This is very interesting if we consider that black immigrants, particular those emigrating from Africa, exhibit the most unfavorable socioeconomic conditions (see for example, Algan at al, 2010, and Dustmann and Theodoropoulos, 2010). Therefore, unobserved cultural and deterrent factors may have a stronger effect for this group.

9 Conclusion

This study investigated the individual relationship between immigration and property crime in England and Wales. Although there is a public sentiment that immigrants are more involved in criminal activities, both the theoretical and the empirical results of this paper lead to different conclusions.

A simple economic model of crime that incorporates immigration has been developed in Section 2. Both this model and other theories developed by sociologists and criminologists illustrated that, even though there are reasons to believe that immigration can be associated with crime, the sign of this association cannot be determined by theory. Therefore, in order to investigate the empirical relationship between immigration and property crime, the Offending, Crime, and Justice Survey of 2003 was employed, a representative national survey of self-reported crime.

Models that account for under-reporting were developed and used, as this is the major concern in crime self-reports. First of all, the empirical analysis showed that under-reporting exists. Despite the fact that under-reporting is not random, meaning that the process to report is covariate-dependent, immigrants' reporting behavior does not differ from natives' one. However, it was explained that the coefficients of the reporting process of the covariate-dependent MisProbit model must be treated with caution, since the reporting process can be also interpreted in a zero-inflation framework. The models indicate that the "mix" of under-reporting and zero-inflation probability, conditional on the set of covariates, is around 70%.

Concerning the crime process, the results of covariate-dependent MisProbit model shows that, after controlling for under-reporting (or zero-inflation) and for basic demographic characteristics, the predicted probability to commit a property crime is about 29%, much higher than the conventional Probit model suggests (about 6%). Most importantly, according to the findings of the covariate-dependent MisProbit, there is a negative but not statistically significant association between immigration and crime (which, however, is significant at 10% if we do not control for ethnicity, or, when we use the unweighted data). Exploiting the extra information of the count form of the property crime variable, and using a NB2-Logit framework, we get to the same conclusion. Natives commit more crimes than immigrants, although this difference is statistically insignificant. The effects of the other covariates on crime are also robust across the binary and the count data models. Moreover, it seems that immigrants are slightly less involved in violent criminal activities as well. Finally, it is quite important that the results of the models used in empirical analysis are not driven by the exclusion restriction of "truthfulness" from the crime process.

Finally, the use of interaction terms offered some interesting insights. Immigrants located in London are considerably less involved in property crime activities than natives. Contrary to that, immigrants in South are more crime-prone than immigrants in London, but not more crime-prone than natives in South (although South immigrants' effect on the probability to commit a crime is higher than all other groups but statistically insignificant). Thus, different socio-economic conditions that immigrants encounter in different locations, and their interactions with the native population, may affect their criminal behavior. Finally, the decomposition of immigrants by ethnic group showed that black immigrants display a considerably lower probabilities of committing a property crime than black natives and white natives, despite the fact that they are the least favored group regarding their economic outcomes.

References

Akins, S., R. G. Rumbaut, and Richard Stansfield (2009), "Immigration, Economic Disadvantage, and Homicide: A Community-level Analysis of Austin, Texas", *Homicide Studies*, 13, 307-314.

Albrecht, H. (1987), "Foreign Minorities and the Criminal Justice System in The Federal Republic of Germany", *The Howard Journal of Criminal Justice*, 26, 272-286.

Albrecht, H. (1997), "Ethnic Minorities, Crime, and Criminal Justice in Germany", in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspective* - Crime and Justice: *A Review of Research*, 21, 31-93. University of Chicago Press.

Algan, Y., C. Dustmann, A. Glitz, and Alan 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.

Armstrong, C. P. (1935), "Some Mental and Social Inadequates", *The Journal of Abnormal and Social Psychology*, 30, 371-383.

Artis, M., M. Ayuso, and M. Guillen (2002), "Detection of Automobile Insurance Fraud with Discrete Choice Models and Misclassified Claims", *Journal of Risk and Insurance*, 69, 325-340.

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.

Bianchi, M., P. Buonanno, and P. Pinotti (2008), "Do Immigrants Cause Crime?", Paris School of Economics, Working Parer, No. 2008 - 05.

Blanchflower, D. G., J. Saleheen, and C. Shadforth (2007), "The Impact of the Recent Migration from Eastern Europe on the UK Economy", IZA Discussion Paper, No. 2615.

Block, M. K., and J. M. Heineke (1975), "A Labor Theoretic Analysis of the Criminal Choice", *American Economic Review*, 65, 314-325.

Bonin, H., A. Constant, K. Tatsiramos, and K. F. Zimmermann (2009), "Native-Migrant Differences in Risk Attitudes", *Applied Economics Letters*, 16, 1581-1586.

Borhas, G. J. (1999), "The Economic Analysis of Immigration", in: Ashenfelter O. and D. Card (ed.) *Handbook of Labor Economics*, 3(28), 1697-1760. Elsevier

Bucher, F. K., and A. M. Piehl (1998a), "Cross-City Evidence on the Relationship between Immigration and Crime", *Journal of Policy Analysis and Management*, 17, 457-493.

Bucher, F. K., and A. M. Piehl (1998a), "Recent Immigrants: Unexpected Implications for Crime and Incarceration", *Industrial and Labor Relations Review*, 51, 654-679.

Bucher, F. K., and A. M. Piehl (2007), "Why are Immigrant' Incarceration Rates so Low? Evidence on Selective Immigration, Deterrence, and Deportation", NBER Working Paper, No. 13229.

Budd, T., C. Sharp, and P. Mayhew (2005), "Offending in England and Wales: First Results from the 2003 Crime and Justice Survey", Home Office Researlch Study 275.

Budd, T., P. Collier, B. Mhlanga, C. Sharp, and G. Weir (2005), "Levels of Self-report Offending and Drug Use Among Offenders: Findings from the Criminality Surveys", Home Office Online Report 18/05.

Burdett K., R. Lagos, and R. Wright (2003), "Crime, Inequality, and Unemployment", *American Economic Review*, 93, 1764-1777.

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

Cameron, A. C., and P.K. Trivedi (2005), *Microeconometrics. Methods and Applications*, New York, Cambridge University Press.

Cameron, S. (1988), "The Economics of Crime Deterrence", *Kyklos*, 41, 301-323.

Caudill, S. B., and F. G. Mixon (2005), "Practitioners' Corner. Analysing Misleading Discrete Responses: A Logit Model Based on Misclassified Data", *Oxford Bulletin of Economics and Statistics*, 67, 105-113.

Chapin, W. D. (1997), "Auslnder Raus? The Empirical Relationship between Immigration and Crime in Germany", *Social Science Quarterly*, 78, 543-557.

Chavez J. M, and E. Griffiths (2009), "Neighborhood Dynamics of Urban Violence: Understanding the Immigration Connection", *Homicide Studies*, 13, 261-273.

Chiu, W. H., and P. Madden (1998), "Burglary and income inequality", *Journal of Public Economics*, 69, 123-141.

Cornwell, C., and W. N. Trumbull (1994), "Estimating the Economic Model of Crime with Panel Data", *Review of Economics and Statistics*, 76, 360-366.

Dustman, C., F. Fabbri, I. Preston and J. Wadworth (1997), "The Distribution of Discrimination in Immigrant Earnings - Evidence from Britain 1974-93", Institute for Fiscal Studies, Working Paper no.97/19.

Dustman, C., F. Fabbri, I. Preston and J. Wadworth (2003), "The Local Labour Market Effects of Immigration in the UK", Home Office Online Report 06/03.

Dustman, C., F. Fabbri and I. Preston (2005), "The Impact of Immigration on the British Labour Market", *Economic Journal*, 113, F324-F341.

Dustman, C., and Nikolaos Theodoropoulos (2010), "Ethnic Minority Immigrants and their Children in Britain", *Oxford Economic Papers*, 62, 209-233.

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

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

Ehrlich, I. (1975), "On the Relation Between Education and Crime", in F. T. Juster (ed): Education, Income, and Human Behavior, 313-338, NBER.

Eide, E. (2000), "Economics of Criminal Behavior", *Encyclopedia of Law and Economics*, 345-389, 8100.

Falaris E. M. (2007), "Misclassification of the Dependent Variable in Binary Choice Models: Evidence from Five Latin American Countries", Department of Economics, University of Delaware, Working Paper No. 2007-05.

Farrington, D. P., B. Gallagher, B. Morley, L. St. Ledger, and D. J. West (1990), "Minimizing Attrition in Longitudinal Research: Methods of Tracing and Securing Cooperation in a 24-year Follow-up Study", in D. Magnusson, and L. Bergman (ed): *Data Quality in Longitudinal Research*, 122-147, Cambridge: Cambridge University Press.

Farrington, D. P. (1986), "Age and Crime", *Crime and Justice*, 7, 189-250.

Farrington, D. P. (2003), "What Has Been Learned from Self-Reports About Criminal Careers and the Causes of Offending?", Report to the Home Office.

Feilzer, M., and R. Hood (2004), "Differences or Discrimination: Minority Ethnic Young People in the Youth Justice System", London: Youth Justice Board.

Feinstein, J. S. (1989), "The Safety Regulation of U.S. Nuclear Power Plants: Violations Inspections, and Abnormal Occurrences," *Journal of Political Economy*, 97, 115-154.

Feinstein, J. S. (1991), "An Econometric Analysis of Income Tax Evasion and Its Detection," *RAND Journal of Economics*, 22, 14-35.

Feldmeyer, B., and D. Steffensmeier (2009), "Immigration Effects on Homicide Offending for Total and Race/Ethnicity-Disaggregated Populations (White, Black, and Latino), *Homicide Studies*, 13, 211-226.

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Volume 1, Edition 3, New York, John Wiley.

Freeman, R. B. (1999), "Economics of Crime", *Handbook in Economics*, 52, 3529-3571.

Gaudry, M. J. I. (1980), "Dogit and Logit Models of Travel Mode Choice in Montreal", *Canadian Journal of Economics*, 13, 268-279.

Gaudry, M. J. I., and M. Dagenais (1979), "The Dogit Model", *Transportation Research*, 13B, 105-111.

Gilpin, N., M. Henty, S. Lemos, J. Portes, and C. Bullen (2006): "The Impact of Free Movement of Workers from Central and Eastern Europe on the UK Labour Market", Department for Work and Pensions, Working Paper No. 29.

Glaeser, E. L., and B. Sacerdote (1999), "Why is There More Crime in Cities?", *Journal of Political Economy*, 107, S225-S258.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996), "Crime and Social Interactions", *Quarterly Journal of Economics*, 111, 507-548.

Graif, C., and R. J. Sampson (2009), "Spatial Heterogeneity in the Effects of Immigration and Diversity on Neighborhood Homicide Rates", *Homicide Studies*, 13, 242-260.

Greene, H. W. (2003), *Econometric Analysis*, 5th Ed., New Jersey, Prentice Hall, Pearson Education International.

Hagan, F. E. (1993), *Research Methods in Criminal Justice and Criminology*, 3rd Edition, New York: Macmillan.

Hagan, J., and A. Palloni (1998), "Immigration and crime in the United States", in J.P. Smith and B. Edmonston (ed): *The immigration debate*, edited by. Washington, D.C. National Academy Press.

Hamlyn, B., C. Maxwell, J. Hales, C. Tait (2003), 2003 Crime & Justice Survey (England and Wales), Technical Report, Home Office.

Hatton, T. J. (2005), "Explaining Trends in UK Immigration", *Journal of Population Economics*, 18, 719-740.

Hatton, T. J., and M. Tani (2005), "Immigration and Inter-regional Mobility in the UK, 1982-2000", *Economic Journal*, 115, F342-F358.

Hausman, J. A., J. Abrevaya, and F. M. Scott-Morton (1996), "Misclassification of the Dependent Variable in a Discrete-Response Setting", mimeo.

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. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement*, 5, 475-492.

Hourwich, I. A. (1912), "Immigration and Crime", *American Journal of Sociology*, 17, 478-490.

Jaeger, D. A., T. Dohmen, A. Falk, D. Huffman, U. Sunde, and H. Bonin (2010), "Direct Evidence on Risk Attitudes and Migration", *The Review of Economics and Statistics*, 92, 684-689.

Junger-Tas, J. (1997), "Ethnic Minorities and Criminal Justice in the Netherlands", in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspectives* - Crime and Justice: *A Review of Research*, 21, 257-307. University of Chicago Press.

Junger-Tas, J., and I. H. Marchall (1999), "The Self-Report Methodology in Crime Research", *Crime & Justice*, 291-367.

Kelly, M. (2000), "Inequality and Crime", *Review of Economics and Statistics*, 82, 530-539.

Kershaw, C., S. Nicholas and A. Walker (2008), "Crime in England and Wales 2007/08", Home Office Statistical Bulletin, 07/08.

Killias, M. (1997), "Immigrants, Crime, and Criminal Justice in Switzerland", in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspectives* - Crime and Justice: *A Review of Research*, 21, 375-405. University of Chicago Press.

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

Lee, M. T., R. Martinez, and R. Rosenfeld (2001), "Does Immigration Increase Homicide? Negative Evidence From Three Border Cities", *The Sociological Quarterly*, 42, 559-580.

Leece, D. (2000), "Household Choice of Fixed Versus Floating Rate Debt: A Binomial Probit Model with Correction for Classification Error", *Oxford Bulletin of Economics and Statistics*, 62, 61-82.

Lemos, S., and J. Portes (2008), "New Labour? The Impact of Migration from Central and Eastern European Countries on the UK Labour Market", IZA Discussion Papers

Lochner, L. (2004), "Education, Work, and Crime: A Human Capital Approach", *International Economic Review*, 45, 811-843.

Lochner, L., and E. Moretti (2001), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports", NBER Working Paper No. 8605.

Lochner, L., and E. Moretti (2004), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports", *American Economic Review*, 94, 155-189.

MacDonald Z. (2001), "Revisiting the Dark Figure. A Microeconometric Analysis of the Underreporting of Property Crime and its Implications," *The British Journal of Criminology*, 41, 127-149.

MacDonald Z. (2002), "Official Crime Statistics: Their Use and Interpretation", *The Economic Journal*, 112, 85-106.

Machin, S., and C. Meghir (2004), "Crime and Economic Incentives", *Journal of Human Resources*, 39, 958-979.

Martens, P. L. (1997), "Immigrants, Crime, and Criminal Justice in Sweden, in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspectives* - Crime and Justice: *A Review of Research*, 21, 183-256. University of Chicago Press.

Martinez R., and M. T. Lee (2000), "On Immigration and Crime", *Criminal Justice*, US Department of Justice, Office of Justice Programs, 1, 486-524.

McBride, D., and C. B. McCoy (1993), "The Drugs-Crime Relationship: An Analytical Framework", *The Prison Journal*, 73, 257-278.

Mears, D. P. (2001), "The Immigration-Crime Nexus: Towards an Analytic Framework for Assessing and Guiding Theory, Research and Policy", *Sociological Perspectives*, 44, 1-19.

Merton, R. K. (1938), "Social Structure and Anomie", *American Sociological Review*, 3, 672-682.

Myers, S. L. (1983), "Estimating the Economic Model of Crime: Employment versus Punishment Effects", *Quarterly Journal of Economics*, 98, 157-166.

Mukhopadhyay, K., and P.K. Trivedi (1995), "Regression Models for Under-recorded Count Data," Paper presented at the Econometric Society 7th World Congress, Tokyo.

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

Ousey, G. G., and C. E. Hubrin (2009), "Exploring the Connection between Immigration and VIolent Crime Rates in U.S Cities, 1980-2000", *Social Problems*, 56, 447-473.

Papadopoulos, G. (2010), "Identifying the Poisson-Logistic Regression Model: Problems, Solutions and Extensions," First chapter of PhD Thesis, Department of Economics, University of Essex.

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

Polczynski O. C., M. K. Laurikkala, L. Huff-Corzine, and J. Corzine (2009), "Immigration and Violent Crime: Citizenship Status and Social Disorganization", *Homicide Studies*, 13, 227-241.

Raphael, S., and R. Winter-Ebmer (2001), "Identifying the Effect of Unemployment on Crime", *Journal of Law and Economics*, 44, 259-283.

Reid, L.W., H.E. Weiss, R.M. Adelman, and, C. Jaret (2005), "The Immigration-Crime Relationship: Evidence across US metropolitan areas", *Social Science Research*, 34, 757-780.

Reilly, B., and R. Witt (1996), "Crime, Deterrence and Unemployment in England and Wales: An Empirical Analysis", *Bulletin of Economic Research*, 48, 137-159.

Sellin, T. (1938), "Culture Conflict and Crime". New York: Social Science Research Council.

Sharp, C., and T. Budd (2005), "Minority Ethnic Groups and Crime: Findings from the Offending, Crime and Justice Survey 2003", 2nd Edition, Home Office Online Report 33/05.

Sjoquist, D. L. (1973), "Property Crime and Economic Behavior: Some Empirical Results", *American Economic Review*, 63, 439-446.

Smith, A. (2006), "Crime Statistics: An Independent Review", Carried out for the Secretary of State for the Home Department.

Smith, D. J. (1997), "Ethnic Origins, Crime, and Criminal Justice in England and Wales", in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspectives* - Crime and Justice: *A Review of Research*, 21, 101-182. University of Chicago Press.

Steffensmeier, D., and E. Allan (1996), "Gender and Crime: Toward a Gendered Theory of Female Offending", *Annual Review of Sociology*, 22, 459-487.

Stowell, I. J. (2007), *Immigration and Crime: The Effects of Immigration on Criminal Behavior*, LFB Scholarly Publishing LLC, New York

Stowell, I. J., and R. Martinez (2009), "Incorporating Ethnic-Specific Measures of Immigration in the Study of Lethal Violence", *Homicide Studies*, 13, 315-324.

Stowell, I. J., S. F. Messner, K. F. McGeever, L. E. Raffalovich (2009), "Immigration and the Recent Violent Crime Drop in the United States: A Pooled, Cross-Sectional Time-Series Analysis of Metropolitan Areas", *Criminology*, 47, 889-928.

Swait, J., and M. Ben-Akiva (1987), "Empirical Test of a Constrained Choice Discrete Model: Mode Choice in Sao Paulo. Brazil", *Transportation Research*, 21B, 103-115.

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.

Thornberry, T. P., and M. D. Krohl (2000), "The Self-Report Method for Measuring Delinquency and Crime", *Criminal Justice*, 4, 33-83.

Tonry, M. (1997), "Ethnicity, Crime, and Immigration", in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspectives* - Crime and Justice: *A Review of Research*, 21, 1-29. University of Chicago Press.

Tournier, P. (1997), "Nationality, Crime, and Criminal ustice in France", in Michael Torny (ed): Ethnicity, Crime, and Immigration. *Comparative and Cross-National Perspectives* - Crime and Justice: *A Review of Research*, 21, 523-5. University of Chicago Press.

Vazsonyi, A. T., and M. Killias (2001), "Immigration and Crime among Youth in Switzerland", *Criminal Justice and Behavior*, 28, 329-366.

Vélez, M. B. (2009) "Contextualizing the Immigration Crime Effect: An Analysis of Homicide in Chicago Neighborhoods", *Homicide Studies, 13, 325-335*

Wadswarth, T. (2010), "Is Immigration Responsible for the Crime Drop? An Assessment of the Influence of Immigration on Changes in Violent Crime Between 1990 and 2000", *Social Science Quarterly*, 91, 531-553.

White, H. R., and D. M. Gorman (2000), "Dynamics of the Drug-Crime Relationship", *Criminal Justice*, US Department of Justice, Office of Justice Programs, 1, 151-218.

Winkelmann, R. (1998), "Count Data Models with Selectivity", *Econometric Reviews*, 17, 339-359.

Winkelmann, R. (2008), *Econometric Analysis of Count Data*, 5th 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.

Witte, A. D. (1980), "Estimating the Economic Model of Crime with Individual Data", *Quarterly Journal of Economics*, 94, 57-84.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MIT Press.

Woplin, K. (1978), "An Economic Analysis of Crime and Punishment in England and Wales, 1894-1967", *Journal of Political Economy*, 86, 815-840.

Yeager, M. G. (1997), "Immigrants and criminality: A review", *Criminal Justice Abstracts*, 29, 143-171.

Appendix A. Zero-Inflated MisProbit

As mentioned in Section 5, it would be a good idea to incorporate a zero inflation probability in the MisProbit model in an attempt to separate under-reporting from zero-inflation. In other words this model will attempt to separate potential criminals from genuine non criminals. In this Appendix this model together with some empirical results are presented. To this end, assume that there is a fraction of people, ξ, that never commit and consequently never report a crime. The remaining fraction of individuals, $1-\xi$, follow the binary choice model with misclassification. The corresponding response tree is presented in figure 5.2. Therefore, the conditional probabilities for the reported crime now become,

$$\Pr(y_i = 1|x_i) = (1-\xi)\left[(1-F(x_i'\beta))\, a_1 + F(x_i'\beta)(1-a_0)\right] \tag{A.1}$$

$$\Pr(y_i = 0|x_i) = \xi + (1-\xi)\left[(1-F(x_i'\beta))\,(1-a_1) + F(x_i'\beta)a_0\right] \tag{A.2}$$

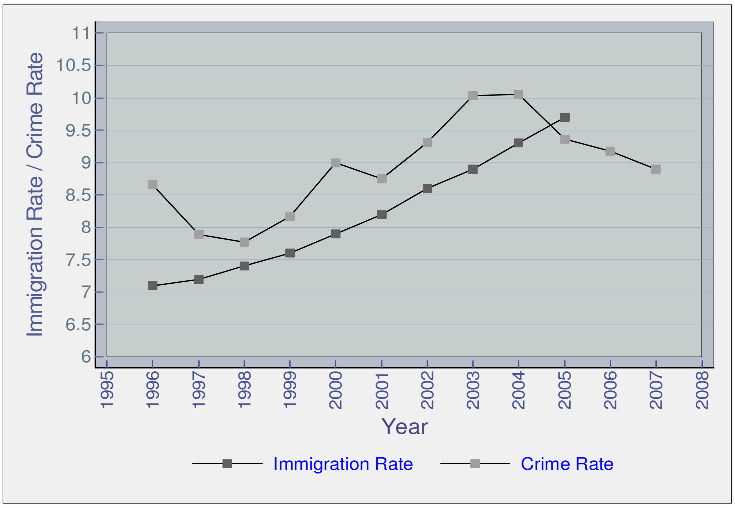
Then, we specify the log-likelihood function as in (5.7) and we find the values of ξ, a_0, a_1, and β that maximize it. In case the probability of zero inflation is given by a Logit model, such as $\xi_i = e^{q_i'u}/(1+e^{q_i'u})$, the log-likelihood takes the following form,

$$\begin{aligned}\ln L(\beta, u, a_0, a_1) = \sum_{i=1}^{n} -\ln(1+e^{q_i'u}) + y_i \ln\left[(1-F(x_i'\beta))a_1 + F(x_i'\beta)(1-a_0)\right] \\ + (1-y_i)\ln[e^{q_i'u} + (1-F(x_i'\beta))(1-a_1) + F(x_i'\beta)a_0].\end{aligned} \tag{A.3}$$

Estimation of this model is quite difficult if probabilities of misclassification and zero-inflation are all covariate-dependent, since with one data set we try to estimate four distinct processes. Instead, given quite large samples, estimation could be feasible if for example, zero-inflation probability is allowed to depend on regressors but one of the misclassification probabilities is considered as constant. In any way, estimation of these models is a hard task, particularly when noisy data such as crime data are used.

The estimation analysis has shown that, although all the coefficients are identified, misclassification and zero-inflation probabilities are not identified if they are considered as constants. Zero-inflation parameter remains unidentified even when under-reporting

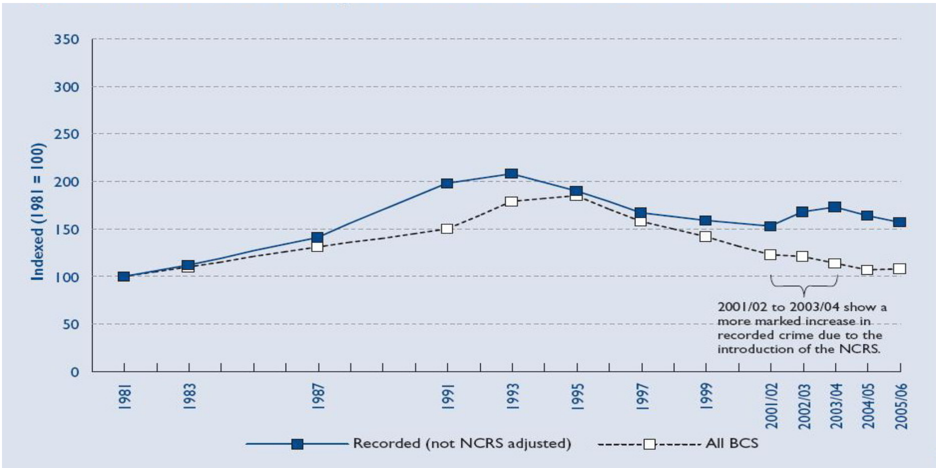
Figure 1.1. Immigration Rates and Crime Rates through time*



The Immigration rate statistics are provided by the OECD Stat. Extracts

The Recorded Crime rate statistics are constructed using data from the Home Office – Research Development Statistics.

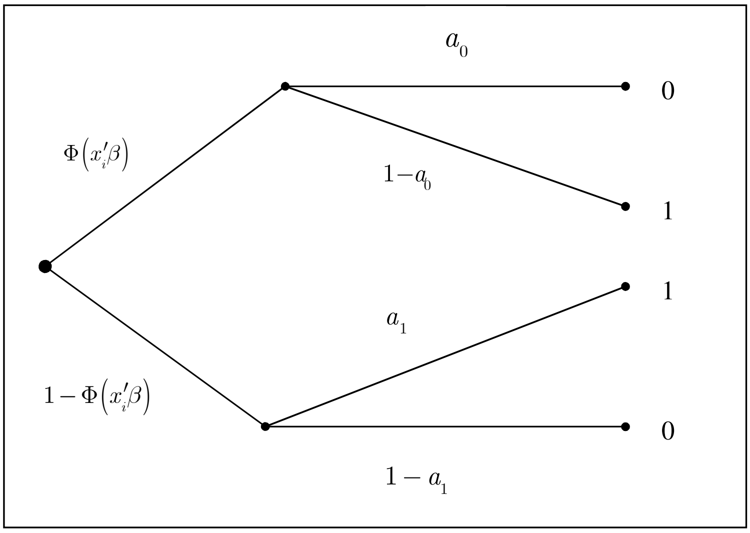
Figure 1.2. Crime Indexes through time: Recorded crime and BCS†



* This graph is constructed without adjustments for the change in recording of crime method happened in April 1998 and the introduction of the National Crime Recording Standard (NCRS) across England and Wales in April 2002. Both changes had the effect of increasing the number of crimes recorded by the police and thus, numbers of recorded crimes are not comparable with previous years. Therefore, the positive tendency for recorded crime between 1998 and 2003 can be considered as a result of these changes, and not as a true increase in crime rates. In figure 2, where the crime data are adjusted for the change in 1998 but not for the introduction of NCRS, it is clear that there is a negative trend up to 2002. The British Crime Survey also coincides with this negative growth rate for crime. It needs to be stressed that criminologists consider the BCS as more reliable than recorded by police crime, since many crimes are not reported to the police, and some reported crimes are not recorded.

† This figure is taken by the independent review of Smith (2006), carried out for the Home Office, page 2.

Figure 5.1. Probit with Misclassification



or

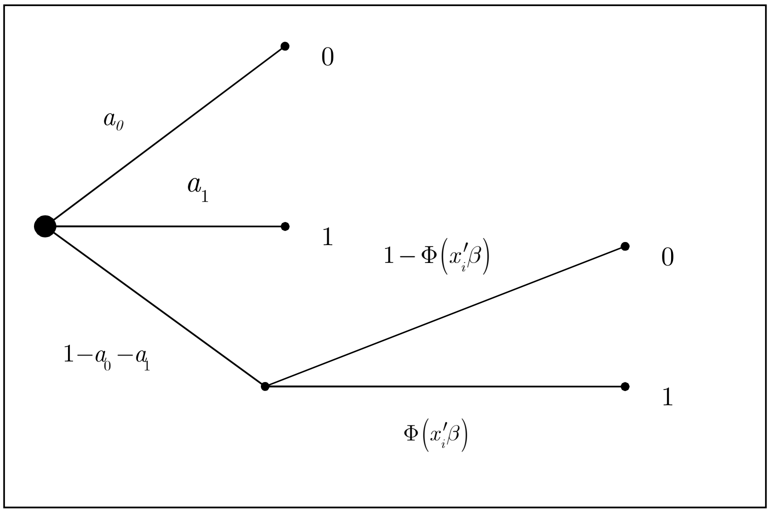


Table 1.1 ESS 2002

Immigrants make country's crime problems worse or better	%
Crime problems made worse	8.3
1	5.8
2	13.5
3	16.5
4	16.9
5	30.8
6	4.0
7	2.2
8	1.3
9	0.3
Crime problems made better	0.3
Total	100

Table 1.2. ISS 1995(n=996) / 2003(n=834)

Immigrants increase crime rates	1995	2003
Agree strongly	7.8	13.6
Agree	18.2	26.3
Neither agree nor disagree	34.9	32.6
Disagree	31.7	24.5
Disagree strongly	7.3	3.1
Total	100	100

Source:ESS(2002),ISS(2003)

Table 1.3. Ordered Probit. Determinants of Natives' Attitudes

Immigrants Increase Crime Rates	Coefficients	Robust Standard Errors
ISS 2003	0.372***	*(0.054)*
Male	0.267***	*(0.055)*
Age	0.008***	*(0.002)*
Income	-0.013**	*(0.006)*
Center	0.049	*(0.076)*
Right	0.368***	*(0.067)*
No party	0.225***	*(0.079)*
Low Education	0.599***	*(0.073)*
Middle Education	0.390***	*(0.066)*
N	1,635	
Log-Likelihood	-2,247.69	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

Table 6.1. Tabulation of Number of Property Crimes

	Frequency	Percent	Cum.
0	10,927	94.17	94.17
1	251	2.16	96.33
2	123	1.06	97.39
3	67	0.58	97.97
4	40	0.34	98.31
5	48	0.41	98.72
6	29	0.25	98.97
7	6	0.05	99.03
8	11	0.09	99.12
9	5	0.04	99.16
10	12	0.1	99.27
11	14	0.12	99.39
12	9	0.08	99.47
13	4	0.03	99.5
14	1	0.01	99.51
15	3	0.03	99.53
16	1	0.01	99.54
17	1	0.01	99.55
18	1	0.01	99.56
19	4	0.03	99.59
20	8	0.07	99.66
21	1	0.01	99.67
22	3	0.03	99.7
23	3	0.03	99.72
24	1	0.01	99.73
25	2	0.02	99.75
27	1	0.01	99.76
28	1	0.01	99.77
30	4	0.03	99.8
33	1	0.01	99.81
34	1	0.01	99.82
35	4	0.03	99.85
36	1	0.01	99.86
40	2	0.02	99.88
41	1	0.01	99.89
50	1	0.01	99.9
54	1	0.01	99.91
55	1	0.01	99.91
56	1	0.01	99.92
57	1	0.01	99.93
60	1	0.01	99.94
73	1	0.01	99.95
100	2	0.02	99.97
113	1	0.01	99.97
168	1	0.01	99.98
194	1	0.01	99.99
225	1	0.01	100
Total	11,604	100	

Table 6.2. Descriptive Statistics

Variables	N			Mean	Weighted Mean			Min	Max
	All	Nat.	Imm.		All	Nat.	Imm.		
<u>Crime Variables</u>									
Any Property Crime last year	11,658	9,647	2,011	0.063	0.055	0.057	0.039	0	1
Any Violent Crime last year	11,667	9,641	2,026	0.072	0.054	0.056	0.036	0	1
Number of Property Crime last year	11,604	9,598	2,006	0.371	0.342 *(5.858)*◊	0.366 *(6.011)*	0.160 *(2.218)*	0	225∅
<u>Independent Variables</u>									
Immigrant	2,069			0.174	0.119∇			0	1
Native	9,853			0.826	0.881				
Age	11,922	9,853	2,069	32.549	36.738 *(17.473)*	36.554 *(17.146)*	38.098 *(19.017)*	10	66
Male	5,755	4,748	1,007	0.483	0.497	0.496	0.505	0	1
Female	6,167	5,105	1,062	0.517	0.503	0.504	0.495		
White	9,284	8,702	582	0.779	0.909	0.956	0.553	0	1
Black	743	291	452	0.062	0.023	0.010	0.120	0	1
Asian	1,116	496	620	0.094	0.045	0.022	0.214	0	1
Other	350	91	259	0.029	0.012	0.003	0.073	0	1
Mixed	429	273	156	0.036	0.012	0.008	0.039	0	1
North	3,249	2,898	351	0.273	0.274	0.288	0.175.	0	1
Midlands	2,822	2,480	342	0.237	0.235	0.246	0.150	0	1
South	3,856	3,352	504	0.323	0.351	0.358	0.298	0	1
London	1,992	1,122	870	0.167	0.139	0.107	0.376	0	1
Truthfulness	11,118	9,271	1,847	0.933	0.942	0.946	0.915	0	1
Other Present	3,768	3,171	597	0.327	0.285	0.288	0.263	0	1

◊ Weighted standard deviations in paretheses.

∅ The max for immigrants is 60 property crimes, while the max for natives is 225 property crimes.

∇ Notice that the weighted mean for the core sample only is 0.091 which is very close to the percentage of immigrants in the UK from other sources. The weighted mean presented here is calculated from the combining sample (core & youth boost & ethnic minorities boost). Although the weights are used to restore representativeness of the sample, these weights are designed to restore representativeness with respect to age and race composition (and also with respect to non respondents). Therefore, it is not surprising to notice a 2.8 percentage points difference.

Table 7.1.a. Probit Estimates

Any ….. in last year	Coefficient	Robust St.Error	Log - Likelihood	N
Property Offence	-0.184**	*(0.089)*	-2,467.26	11,658
Violent Offence	-0.209***	*(0.082)*	-2,440.57	11,667
Drugs related Offence	-0.091	*(0.144)*	-680.19	11,866
Vehicle Theft	-0.137	*(0.255)*	-358.43	11,873
Criminal Damage	-0.468***	*(0.134)*	-693.62	11,858
Burglary	-0.485**	*(0.210)*	-131.93	11,870
Robbery	-0.231	*(0.277)*	-31.87	11,897
Other Theft	-0.149	*(0.091)*	-2188.28	11,713
Assault	-0.210***	*(0.082)*	-2438.79	11,676

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.1.b. Negative Binomial Estimates

Number of ….. in last year	Coefficient	Robust St.Error	Alpha	Log - Likelihood	N
Property Offences	-0.825**	*(0.351)*	57.46***	-2,400.98	11,604
Violent Offences	-1.062***	*(0.236)*	50.02***	-2,397.71	11,640
Drugs related Offences	-0.144	*(0.645)*	406.34***	-675.27	11,862
Vehicle Thefts	-1.792**	*(0.722)*	542.04***	-289.88	11,869
Criminal Damages	-2.451***	*(0.494)*	182.86***	-571.78	11,856
Burglaries	-3.035***	*(0.932)*	2,123.43***	-115.21	11,869
Robberies	-2.373**	*(1.105)*	7,695.16***	-25.65	11,897
Other Thefts	-0.664*	*(0.366)*	64.41***	-2,157.15	11,695
Assaults	-1.060***	*(0.237)*	50.05***	-2,393.85	11,649

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.2.a. Probit Estimates

Probit	(1)		(2)		(3)	
	Coefficient	Robust S.E	Coefficient	Robust S.E	Coefficient	Robust S.E
Probability of Committing a Property Offence in Last Year						
Constant	-1.204***	*(0.127)*	-1.454***	*(0.143)*	-1.134***	*(0.129)*
Immigrant	-0.127	*(0.102)*	-0.025	*(0.111)*	-0.014	*(0.109)*
Age	-.0178**	*(0.007)*	-0.018***	*(0.007)*	-0.018**	*(0.007)*
Age^2	0.0001	*(0.0001)*	0.0001	*(0.0001)*	0.0001	*(0.0001)*
Male	0.362***	*(0.049)*	0.362***	*(0.049)*	0.364***	*(0.049)*
White			0.331***	*(0.077)*		
Black					-0.201*	*(0.119)*
Asian & Other					-0.501***	*(0.101)*
Mixed					-0.016	*(0.115)*
Region South	0.089	*(0.082)*	0.028	*(0.082)*	0.034	*(0.083)*
Region Midlands	0.061	*(0.084)*	0.007	*(0.085)*	0.012	*(0.086)*
Region North	0.066	*(0.086)*	0.009	*(0.085)*	0.015	*(0.086)*
Sample Size	11,658		11,658		11,658	
Log Likelihood	-1,452.88		-1,447.94		-1,445.82	
Predicted.Prob.Crime	0.064		0.061		0.062	

Table 7.2.b. Negative Binomial 2 Estimates

NegBin2	(1)		(2)		(3)	
	Coefficient	R.S.E	Coefficient	R.S.E	Coefficient	R.S.E
Expected number of Property Offences in Last Year						
Constant	-1.299***	*(0.505)*	-1.897***	*(0.529)*	-1.065***	*(0.484)*
Immigrant	-0.441	*(0.356)*	-0.309	*(0.352)*	-0.232	*(0.353)*
Age	0.001	*(0.031)*	0.002	*(0.030)*	0.003	*(0.030)*
Age^2	-0.001*	*(0.0004)*	-0.001*	*(0.0004)*	-0.001*	*(0.0004)*
Male	0.669**	*(0.278)*	0.672**	*(0.275)*	0.663**	*(0.276)*
White			0.788***	*(0.268)*		
Black					-0.736**	*(0.328)*
Asian & Other					-1.337***	*(0.366)*
Mixed					0.112	*(0.435)*
Region South	0.504*	*(0.284)*	0.305	*(0.279)*	0.244	*(0.283)*
Region Midlands	0.593**	*(0.272)*	0.441	*(0.269)*	0.370	*(0.274)*
Region North	1.256**	*(0.501)*	1.065**	*(0.507)*	1.008**	*(0.503)*
Sample Size	11,604		11,604		11,604	
Log Likelihood	-2,346.791		-2,344.4531		-2,342.7878	
Alpha	47.06***		46.66***		46.33***	
Predicted.Num.Crimes	0.407		0.388		0.388	

Robust standard errors are in parentheses.

(***) denotes 1%, (**) denotes 5%, and (*) denotes 10% significance level

Table 7.4. Covariate-Dependent Probit allowing for Misclassification

Mis.Probit2	(1)		(2)		(3)	
	Coefficient	Robust S.E	Coefficient	Robust S.E	Coefficient	Robust S.E
Probability of Committing a Property Offence in Last Year						
Constant	2.597***	*(0.895)*	2.335***	*(0.896)*	2.833***	*(0.926)*
Immigrant	-0.431*	*(0.251)*	-0.254	*(0.273)*	-0.277	*(0.275)*
Age	-0.248***	*(0.045)*	-0.259***	*(0.046)*	-0.255***	*(0.047)*
Age^2	0.003***	*(0.001)*	0.003***	*(0.001)*	0.003***	*(0.001)*
Male	0.439***	*(0.156)*	0.440***	*(0.156)*	0.429***	*(0.156)*
White			0.542*	*(0.293)*		
Black					-0.559*	*(0.323)*
Asian and Other					-0.638	*(0.430)*
Mixed					0.004	*(0.391)*
Region South	0.300	*(0.219)*	0.273	*(0.231)*	0.272	*(0.227)*
Region Midlands	0.151	*(0.209)*	0.124	*(0.228)*	0.124	*(0.223)*
Region North	0.501*	*(0.282)*	0.496	*(0.318)*	0.498	*(0.317)*
Prob of Misclassification of One as Zero (Under-reporting or Zero-Inflation)						
Constant	2.196***	*(0.522)*	2.267***	*(0.507)*	2.072***	*(0.516)*
Immigrant	-0.391	*(0.371)*	-0.410	*(0.346)*	-0.437	*(0.359)*
Age	-0.205**	*(0.081)*	-0.195***	*(0.074)*	-0.193***	*(0.074)*
Age^2	0.003**	*(0.001)*	0.003***	*(0.001)*	0.003***	*(0.001)*
Male	-0.260*	*(0.144)*	-0.277*	*(0.142)*	-0.285***	*(0.142)*
White			-0.198	*(0.224)*		
Black					-0.264	*(0.447)*
Asian and Other					0.467	*(0.327)*
Mixed					0.104	*(0.281)*
Region South	0.201	*(0.228)*	0.262	*(0.224)*	0.252	*(0.221)*
Region Midlands	0.151	*(0.215)*	0.224	*(0.215)*	0.211	*(0.212)*
Region North	0.432**	*(0.211)*	0.490**	*(0.212)*	0.480**	*(0.207)*
Truthfulness	0.876***	*(0.319)*	0.868***	*(0.256)*	0.854***	*(0.255)*
Prob of Misclassification of Zero as One (Over-reporting) +						
Constant	-2.244***	*(0.292)*	-2.234***	*(0.204)*	-2.256***	*(0.210)*
Sample Size	11,658		11,658		11,658	
Log Likelihood	-1,427.97		-1,422.05		-1,419.99	
Predicted.Prob.Crime	0.285		0.292		0.295	
Predicted.Prob.Under	0.681		0.709		0.709	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

+ Which corresponds to probability of misclassification of zero as one of $\Phi(-2.234)=0.0127$

Table 7.5. Negative Binomial 2 Models

	NB2		NB2-Logit		ZI-NB2-Logit	
	Coefficient	R.S.E	Coefficient	R.S.E	Coefficient	R.S.E
Expected number of Property Offences in Last Year						
Constant	-1.897***	*(0.529)*	8.633**	*(3.671)*	10.950**	*(5.288)*
Immigrant	-0.309	*(0.352)*	-0.617	*(0.636)*	-0.757	*(0.721)*
Age	0.002	*(0.030)*	-0.677***	*(0.224)*	- 0.711**	*(0.308)*
Age^2	-0.001*	*(0.000)*	0.009***	*(0.003)*	0.010**	*(0.005)*
Male	0.672**	*(0.275)*	1.457***	*(0.443)*	0.657	*(0.498)*
White	0.788***	*(0.268)*	0.196	*(0.611)*	-0.939	*(0.706)*
Region South	0.305	*(0.279)*	1.024	*(0.696)*	1.504**	*(0.694)*
Region Midlands	0.441	*(0.269)*	0.247	*(0.561)*	0.501	*(0.682)*
Region North	1.065 **	*(0.507)*	1.676**	*(0.662)*	2.733***	*(0.949)*
Probability of Reporting a Committed Crime						
Constant			-12.253***	*(3.510)*	-13.285***	*(4.520)*
Immigrant			0.451	*(1.043)*	0.130	*(0.869)*
Age			1.008***	*(0.206)*	1.058***	*(0.230)*
Age^2			-0.015***	*(0.003)*	-0.015***	*(0.004)*
Male			-1.333*	*(0.687)*	-0.406	*(0.625)*
White			0.638	*(1.062)*	1.053	*(0.994)*
Region South			-1.508	*(0.937)*	-1.890**	*(0.771)*
Region Midlands			0.255	*(0.954)*	-0.480	*(0.878)*
Region North			-1.628	*(1.027)*	-2.800***	*(1.010)*
Truthfulness			-1.237***	*(0.475)*	-1.963***	*(0.468)*
Probability of Zero-Inflation						
Constant					-1.212	*(1.133)*
Immigrant					-0.533	*(0.549)*
Age					0.138**	*(0.060)*
Age^2					-0.001	*(0.001)*
Male					-0.960***	*(0.259)*
White					-1.634***	*(0.387)*
Region South					0.490	*(0.480)*
Region Midlands					0.136	*(0.561)*
Region North					0.820*	*(0.479)*
Sample Size	11,604		11,604		11,604	
Log Likelihood	-2,344.45		-2,313.99		-2,258.49	
alpha	46.66***		41.64***		15.474***	
Pred.Pr. of Reporting			0.430		0.376	
Pred.Pr. of ZI					0.617	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.6. Weighted Property Crime *versus* Weighted Violent Crime *versus* Unweighted Property Crime

Mis.Probit2	Weighted Property Crime		Weighted Violent Crime		Unweighted Property Crime	
	Coefficient	Robust S.E	Coefficient	Robust S.E	Coefficient	Robust S.E
Probability of Committing an Offence in Last Year						
Constant	2.335***	*(0.896)*	2.778**	*(0.925)*	2.576***	*(0.569)*
Immigrant	-0.254	*(0.273)*	-0.241	*(0.341)*	-0.493**	*(0.210)*
Age	-0.259***	*(0.046)*	-0.315***	*(0.056)*	-0.270***	*(0.035)*
Age^2	0.003***	*(0.001)*	0.004***	*(0.001)*	0.004***	*(0.001)*
Male	0.440***	*(0.156)*	0.428***	*(0.102)*	0.464***	*(0.125)*
White	0.542*	*(0.293)*	0.447***	*(0.172)*	0.487**	*(0.189)*
Region South	0.273	*(0.231)*	-0.063	*(0.241)*	0.050	*(0.162)*
Region Midlands	0.124	*(0.228)*	-0.085	*(0.242)*	-0.032	*(0.177)*
Region North	0.496	*(0.318)*	-0.140	*(0.246)*	0.103	*(0.201)*
Prob of Misclassification of One as Zero (Under-reporting or Zero-Inflation)						
Constant	2.267***	*(0.507)*	6.174***	*(1.250)*	2.509***	*(0.526)*
Immigrant	-0.410	*(0.346)*	-0.553	*(0.416)*	-0.209	*(0.262)*
Age	-0.195***	*(0.074)*	-0.646***	*(0.154)*	-0.261***	*(0.066)*
Age^2	0.003***	*(0.001)*	0.012***	*(0.003)*	0.004***	*(0.001)*
Male	-0.277*	*(0.142)*	-0.229	*(0.149)*	-0.179	*(0.130)*
White	-0.198	*(0.224)*	-0.057	*(0.264)*	-0.074	*(0.211)*
Region South	0.262	*(0.224)*	-0.237	*(0.310)*	0.151	*(0.178)*
Region Midlands	0.224	*(0.215)*	-0.063	*(0.309)*	0.222	*(0.184)*
Region North	0.490**	*(0.212)*	-0.256	*(0.316)*	0.299	*(0.192)*
Truthfulness	0.868***	*(0.256)*	0.876***	*(0.253)*	1.181***	*(0.197)*
Prob of Misclassification of Zero as One (Over-reporting)						
Constant	-2.234***	*(0.204)*	-2.081***	*(0.088)*	-2.175***	*(0.121)*
Sample Size	11,658		11,667		11,658	
Log Likelihood	-1,422.05		-1,303.74		-2,441.32	
Predicted.Prob.Crime	0.292		0.223		0.260	
Predicted.Prob.Under	0.709		0.514		0.650	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.7.a. Truthfulness *versus* No Exclusion *versus* Other Present, for Property Crime

Mis.Probit2	(1) Property Crime (Truthfulness)		(2) Property Crime (No Exclusion)		(3) Property Crime (Other Present)	
	Coefficient	Rob. S.E	Coefficient	Rob. S.E	Coefficient	Rob. S.E
Probability of Committing an Offence in Last Year						
Constant	2.335***	*(0.896)*	2.298	*(1.558)*	0.985	*(5.056)*
Immigrant	-0.254	*(0.273)*	-0.232	*(0.276)*	-0.143	*(0.528)*
Age	-0.259***	*(0.046)*	-0.242***	*(0.083)*	-0.170	*(0.285)*
Age^2	0.003***	*(0.001)*	0.003***	*(0.001)*	0.002	*(0.004)*
Male	0.440***	*(0.156)*	0.401***	*(0.149)*	0.382***	*(0.122)*
White	0.542*	*(0.293)*	0.369	*(0.297)*	0.369*	*(0.219)*
Region South	0.273	*(0.231)*	0.175	*(0.247)*	0.085	*(0.456)*
Region Midlands	0.124	*(0.228)*	0.066	*(0.203)*	0.022	*(0.212)*
Region North	0.496	*(0.318)*	0.379	*(0.297)*	0.369	*(0.219)*
Prob of Misclassification of One as Zero (Under-reporting or Zero-Inflation)						
Constant	2.267***	*(0.507)*	3.070***	*(0.715)*	2.906**	*(1.179)*
Immigrant	-0.410	*(0.346)*	-0.343	*(0.305)*	-0.387	*(0.399)*
Age	-0.195***	*(0.074)*	-0.192*	*(0.107)*	-0.244	*(0.367)*
Age^2	0.003***	*(0.001)*	0.003**	*(0.001)*	0.004	*(0.005)*
Male	-0.277*	*(0.142)*	-0.239	*(0.169)*	-0.168	*(0.394)*
White	-0.198	*(0.224)*	-0.239	*(0.240)*	-0.103	*(0.760)*
Region South	0.262	*(0.224)*	0.200	*(0.194)*	0.188	*(0.558)*
Region Midlands	0.224	*(0.215)*	0.126	*(0.208)*	0.086	*(0.355)*
Region North	0.490**	*(0.212)*	0.438**	*(0.181)*	0.475*	*(0.266)*
Truthfulness	0.868***	*(0.256)*				
Other Present					0.294	*(0.505)*
Prob of Misclassification of Zero as One (Over-reporting)						
Constant	-2.234***	*(0.204)*	-2.552***	*(0.709)*	-2.941	*(2.019)*
Log Likelihood	-1,422.05		-1,432.94		-1,430.67	
Predicted.Prob.Crime	0.292		0.291		0.170	
Predicted.Prob.Under	0.709		0.684		0.475	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.7.b. Truthfulness *versus* No Exclusion *versus* Other Present, for Violent Crime

Mis.Probit2	(1) Violent Crime (Truthfulness)		(2) Violent Crime (No Exclusion)		(3) Violent Crime (Other Present)	
	Coefficient	Rob. S.E	Coefficient	Rob. S.E	Coefficient	Rob. S.E
Probability of Committing an Offence in Last Year						
Constant	2.778**	*(0.925)*	2.470***	*(0.677)*	2.354***	*(0.645)*
Immigrant	-0.241	*(0.341)*	-0.177	*(0.271)*	-0.169	*(0.259)*
Age	-0.315***	*(0.056)*	-0.296***	*(0.044)*	-0.291***	*(0.040)*
Age^2	0.004***	*(0.001)*	0.004***	*(0.001)*	0.004***	*(0.001)*
Male	0.428***	*(0.102)*	0.410***	*(0.093)*	0.419***	*(0.092)*
White	0.447***	*(0.172)*	0.429***	*(0.158)*	0.447***	*(0.172)*
Region South	-0.063	*(0.241)*	-0.050	*(0.213)*	-0.042	*(0.212)*
Region Midlands	-0.085	*(0.242)*	-0.085	*(0.216)*	-0.074	*(0.210)*
Region North	-0.140	*(0.246)*	-0.128	*(0.158)*	-0.117	*(0.211)*
Prob of Misclassification of One as Zero (Under-reporting or Zero-Inflation)						
Constant	6.174***	*(1.250)*	7.666***	*(0.927)*	7.509***	*(0.995)*
Immigrant	-0.553	*(0.416)*	-0.502	*(0.423)*	-0.503	*(0.440)*
Age	-0.646***	*(0.154)*	-0.723***	*(0.110)*	-0.742***	*(0.135)*
Age^2	0.012***	*(0.003)*	0.014***	*(0.002)*	0.014***	*(0.004)*
Male	-0.229	*(0.149)*	-0.283*	*(0.150)*	-0.276*	*(0.157)*
White	-0.057	*(0.264)*	-0.060	*(0.266)*	-0.071	*(0.269)*
Region South	-0.237	*(0.310)*	-0.186	*(0.319)*	-0.165	*(0.337)*
Region Midlands	-0.063	*(0.309)*	-0.050	*(0.321)*	-0.020	*(0.343)*
Region North	-0.256	*(0.316)*	-0.259	*(0.327)*	-0.252	*(0.330)*
Truthfulness	0.876***	*(0.253)*				
Other Present					0.350***	*(0.160)*
Prob of Misclassification of Zero as One (Over-reporting)						
Constant	-2.081***	*(0.088)*	-2.108***	*(0.085)*	-2.095***	*(0.085)*
Log Likelihood	-1,303.74		-1,307.16		-1,305.09	
Predicted.Prob.Crime	0.223		0.207		0.207	
Predicted.Prob.Under	0.514		0.470		0.483	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.7.c NegBin2-Logit – Truthfulness *versus* Other Present

NegBin2-Logit	(1) Truthfulness		(2) Other Present – A –		(3) Other Present – B –	
	Coefficient	R.S.E	Coefficient	R.S.E	Coefficient	R.S.E
Expected number of Property Offences in Last Year						
Constant	8.633**	*(3.671)*	7.572**	*(3.009)*	-5.849***	*(1.538)*
Immigrant	-0.617	*(0.636)*	-0.475	*(0.595)*	-0.306	*(0.640)*
Age	-0.677***	*(0.224)*	-0.619***	*(0.180)*	0.422***	*(0.145)*
Age^2	0.009***	*(0.003)*	0.009***	*(0.003)*	-0.007***	*(0.002)*
Male	1.457***	*(0.443)*	1.462***	*(0.394)*	0.125	*(0.477)*
White	0.196	*(0.611)*	0.175	*(0.662)*	0.869	*(0.611)*
Region South	1.024	*(0.696)*	0.860	*(0.601)*	-0.270**	*(0.507)*
Region Midlands	0.247	*(0.561)*	0.161	*(0.558)*	0.694	*(0.627)*
Region North	1.676**	*(0.662)*	1.588**	*(0.711)*	-0.201	*(0.575)*
Probability of Reporting a Committed Crime						
Constant	-12.253***	*(3.510)*	-11.470***	*(3.664)*	11.870***	*(2.799)*
Immigrant	0.451	*(1.043)*	0.189	*(0.972)*	-0.094	*(0.931)*
Age	1.008***	*(0.206)*	0.910***	*(0.208)*	-0.939***	*(0.182)*
Age^2	-0.015***	*(0.003)*	-0.014***	*(0.003)*	0.014***	*(0.003)*
Male	-1.333*	*(0.687)*	-1.325**	*(0.659)*	1.139	*(0.708)*
White	0.638	*(1.062)*	0.734	*(1.191)*	-0.525	*(0.974)*
Region South	-1.508	*(0.937)*	-1.024	*(0.902)*	0.992	*(0.771)*
Region Midlands	0.255	*(0.954)*	0.519	*(0.995)*	-0.416	*(0.878)*
Region North	-1.628	*(1.027)*	-1.373	*(1.164)*	1.903	*(0.845)*
Truthfulness	-1.237***	*(0.475)*				
Other Present			-0.677	*(0.533)*	-0.600*	*(0.349)*
Sample Size	11,604		11,604		11,604	
Log Likelihood	-2,313.99		-2,314.29		-2,314.61	
alpha	41.64***		41.94***		41.99***	
Pred.Pr. of Reporting	0.430		0.448		0.464	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table 7.8.a. Interaction Terms

Covariate Dependent MisProbit	(1) Ethnicity		(2) Regions	
	Coef	R.S.E	Coef	R.S.E
Constant	2.777***	*(0.898)*	1.997**	*(0.838)*
Immigrant	-0.196	*(0.287)*	-0.917***	*(0.352)*
Age	-0.252***	*(0.046)*	-0.248***	*(0.036)*
Age^2	0.003***	*(0.001)*	0.003***	*(0.001)*
Male	0.428***	*(0.158)*	0.521***	*(0.140)*
White			0.571	*(0.372)*
Black	0.013	*(0.367)*		
Asian and Other	-0.617	*(0.457)*		
Mixed	0.372	*(0.541)*		
Region South	0.275	*(0.237)*	0.009	*(0.217)*
Region Midlands	0.139	*(0.228)*	-0.051	*(0.226)*
Region North	0.489	*(0.316)*	0.301	*(0.289)*
Immigrant*Black	-0.934	*(0.583)*		
Immigr*Asian&Other	0.054	*(0.730)*		
Immigrant*Mixed	-0.615	*(0.919)*		
Immigrant*South			1.483***	*(0.570)*
Immigrant*Midlands			0.690	*(0.556)*
Immigrant*North			0.240	*(0.539)*
Sample Size	11,658		11,658	
Log Likelihood	-1,418.40		-1,413.96	
Predicted.Prob.Crime	0.297		0.243	

Table 7.8.b. Interaction Terms (Specification (1) Continue)[⊕]

Covariate Dependent MisProbit	(1) Ethnicity	
Immigrant*Asian&Other	0.359	*(0.709)*
Immigrant*Mixed	0.678	*(0.780)*
Immigrant*White	0.921*	*(0.475)*
Native*Black	1.131**	*(0.522)*
Native*Asian&Other	0.501	*(0.585)*
Native*Mixed	1.490**	*(0.669)*
Native*White	1.117***	*(0.404)*

Table 7.8.c. Interaction Terms (Specification (2) Continue)

Covariate Dependent MisProbit	(2) Regions Base Group: Immigrant & London	
Immigrant*London		
Immigrant*South	1.492***	*(0.484)*
Immigrant*Midlands	0.639	*(0.504)*
Immigrant*North	0.541	*(0.459)*
Native*London	0.917***	*(0.352)*
Native*South	0.927***	*(0.308)*
Native*Midlands	0.867***	*(0.320)*
Native*North	0.541***	*(0.459)*

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

[⊕] This model is exactly the same with the one presented in Table 7.7, apart from the way we define the variables associated with the interaction terms. Thus, all the other coefficients are exactly the same with specification (1) of Table 7.7 and therefore, not presented here. The same holds for the second specification depends on covariates. However, if zero-inflation depends on regressors, both ZI-MisProbit model with constant misclassification and covariate dependent ZI-MisProbit model, behave better. First, the case of constant misclassification is presented, followed by the covariate dependent ZI-MisProbit.

Constant Misclassification

Although under-reporting seems to be covariate dependent, here the results of a model of constant under-reporting is presented. Naturally, zero-inflation should depend on the same vector of regressors. Here, truthfulness, which is assumed to affect the zero-inflation probability, is used to facilitate the optimization procedure. However, an extra exclusion was required (here in the form of not including age squared in the reporting process), even though there was no specific reason for this. Otherwise, misclassification probabilities were forced to be negative, and therefore, not helping the separation of zero-inflation from under-reporting. This is also the reason why this model was not presented in the main results analysis, as the behavior of this model was not trustworthy, perhaps because of the combination of noisy data and complicated models. Nevertheless, a few results which will be presented in this Appendix indicate that this model "works", as it can potentially separate zero-inflation from under-reporting.

These results are presented in Table A.1. In the first column results of the MisProbit model with constant misclassification are presented, whereas the ZI-MisProbit results are presented in column 2. It is very interesting that, after controlling for zero-inflation, the probability of under-reporting is predicted to be 0.48, almost 33 percentage points lower than the model in column 1. The predicted probability of zero-inflation is 39.3%. Thus, this model says that almost 40% of people never commit crimes and consequently they do not report any, and from the rest of them, 48% report no crimes even though they have committed at least one.[95] Furthermore, we can notice that all coefficient of the zero-inflation process are statistically insignificant. We finally see that both the predicted probability of committing a crime and the probability of over-reporting is almost the same

[95] Apart from the case where age squared was included in both processes, all other specifications show that given the set of controls, zero-inflation probability is around 40% and probability of under-reporting is around 35%.

across the two models.

Covariate-Dependent Under-reporting

As we have seen from the main results, the mix of under-reporting and zero-inflation is covariate-dependent. In the previous part only zero-inflation was allowed to depend on covariates. In this part we will have a look at the model where both processes are covariate-dependent. Misclassifying a zero as one will still be considered as constant. Before proceeding to the results, we should say that with one data set we try to identify the parameters of three different processes. In addition to that, we must be cautious not to misspecify the model by excluding variables that must be included. For example, in the previous model I did not include age squared in the zero-inflation process without a special reason. Therefore, this model is too demanding to produce reliable estimates with such noisy data. However, some results are presented in Table A.2, which show that this model also "works".

Once more, the first column replicates the second specification of Table 7.4. First of all, we notice that all parameters of this model are identified. It can be said that, this is a "better specified" model, since there is only one exclusion restriction from both zero-inflation and under-reporting process. This role, as before, is played by the dummy, "truthfulness". On the other hand, as mentioned before, this is a too complicated model relative to the quality of data and trustworthiness is a question here.

Unfortunately, the results of this model do not coincide with ZI-MisProbit with constant misclassification. In this model, we see that the predicted probability of crime is around 51%, which says that from potential criminals (since we have separated the genuine non criminals) almost half of them commit at least one property crime. Moreover, we can also notice that the predicted probability of zero-inflation is just around 10%, which is much smaller than the predictions of the previous model. Nonetheless, there is one common finding across these two. In both models, apart from truthfulness which seems to be significant in the current model, the independent variables do not seem to affect the probability of being genuine non-criminal. However, the values of their corresponding coefficients differ a lot. For some of these variables even the direction of the effect is the opposite one.

Comparing the first column with the second, there are a few things that merit some discussion. Although we would expect the predicted probability to be lower, the ZI-MisProbit gives almost the same probability of under-reporting and an extra zero-inflation probability of 8%. Thus, this model, at least for the data of this study, does not seem to separate zero-inflation from under-reporting. Regarding the immigrant coefficient, in contrast with column 1, the ZI-MisProbit model says that being an immigrant increases the probability of crime, but the coefficient is statistically insignificant. Although the rest of the coefficients in the crime process follow the same direction as the coefficients of the first column, they are very different in terms magnitude. The rest of the coefficients of the two processes are not discussed further since this model is just presented to show that if the generating data process follows a ZI-MisProbit model, given a richer data set, there might be gains from using it.

Appendix B. NB1-Logit, Generalized-NB-Logit, ZI-Poisson-Logit, and ZI-NB2-Logit

In subsection 5.2 we discussed that identification of the NB2-Logit model requires exactly the same conditions established for the Poisson-Logit model (exclusion restriction on count process or sign restrictions on reporting process). A model that is identified even without the restrictions described above is the Negative Binomial 1-Logit (NB1-Logit), which is obtained if we assume that α_i depends on regressors in the following manner, $\alpha_i = \theta/\lambda_i$ (see, Papadopoulos, 2010a, and Papadopoulos and Santos Silva, 2008). According to this form of variance of ϵ_i, the variance of y_i changes to $\omega_i = \mu_i + \theta\lambda_i\Lambda_i^2$. It should be noted that identification of the conditional mean is easier only because we impose more structure on the variance. Hence, if the variance form of α_i is misspecified, the estimates of θ will be in general inconsistent.

Instead of assuming the form of the variance, we can specify a generalization of it as $\omega_i = \mu_i + \theta\lambda_i^{2-c}\Lambda_i^2$, where c is an extra parameter to be estimated. In case $c = 0$, a NB2-Logit is obtained, whereas in case $c = 1$, a NB1-Logit is obtained. Therefore, identification becomes "weaker" as c gets closer to 0. According to this general parameterization of the variance the following log-likelihood arises,

$$\ln L(\theta, c, \beta, \gamma) = \sum_{i=1}^{n} \ln \left(\Gamma(y_i + \theta^{-1}\lambda_i^c)/\Gamma(y_i+1)\Gamma(\theta^{-1}\lambda_i^c) \right) - \\ (\theta^{-1}\lambda_i^c + y_i)\ln(1+\theta\lambda_i^{1-c}\Lambda_i) + y_i(\ln\lambda_i^{1-c} + \ln\Lambda_i + \ln\theta) \quad \text{(A.4)}$$

Similarly to the models for binary choice, models for count data with under-reporting can also be generalized to take into account zero-inflation. First, the Zero-Inflation-Poisson-Logit (ZIP-Logit) specification is presented. A construction of a response tree similar to (5.2) is helpful to derive the conditional probabilities of interest. As before, there is a fraction of people, ξ, that never commit and consequently never report crimes. The remaining fraction of individuals, $(1-\xi)$, can either commit or not commit crimes, but their responses are subject to under-reporting, meaning that they follow the Poisson-Logit model. Therefore, zeroes come from zero-inflation, or, from the Poisson-Logit mixture distribution. That is, zeroes because of under-reporting, or zeros because of the choice not to commit crimes. The response tree is presented in figure 5.3. In this case, ξ cannot distinguish between zeros because of never committing crimes (zero inflation) and always reporting zeroes (total under-reporting), which was the case in the binary choice model.

In this tree, $e^{-\mu_i}$, is the probability of zero from the Poisson-Logit model, and $\mu_i = \lambda_i\Lambda_i$. According to this model, the conditional probabilities of zero and positives are the following,

$$\begin{aligned} \Pr(y_i = 0|x_i) &= \xi + (1-\xi)e^{-\mu_i}, \\ \Pr(y_i > 0|x_i) &= (1-\xi)\frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}. \end{aligned} \quad \text{(A.5)}$$

The mean of this model is given by $\nu_i = (1-\xi)\mu_i$ and variance $\omega_i = (1-\xi)(1+\xi\mu_i)\mu_i$, so that there is overdispersion. If ξ is a constant, the log-likelihood is given by,

$$\ln L(u, \beta, \gamma) = \sum_{y=0} \ln \left(\xi + (1-\xi)e^{-\mu_i} \right) + \sum_{y>0} \ln \left((1-\xi)\frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \right) \quad \text{(A.6)}$$

However, ξ can also depend on covariates, so that we try to model what are the characteristics that lead people never committing crimes or always reporting no crimes. If ξ is a Logit, with $\xi_i = e^{q_i'u}/(1+e^{q_i'u})$, then we can write the log-likelihood function of the ZIP-Logit as,

$$\begin{aligned}\ln L(u,\beta,\gamma) = -\sum_{i=1}^{n} \ln\left(1+e^{q_i'u}\right) + \sum_{y=0}\left(e^{q_i'u} - e^{-\mu_i}\right) \\ + \sum_{y>0}(-\mu_i + y_i \ln \mu_i - \ln(y_i!))\end{aligned} \tag{A.7}$$

Identification of this model requires the same assumptions established for the Poisson-Logit, so that exclusion restrictions in the Poisson part, or sign restrictions on the Logit part are required.

Reformulation as a Zero-Inflation-NB2-Logit (ZI-NB2-Logit) model is straightforward. The probability of an observed zero outcome from the NB2-Logit is now given by $(1 + \alpha\mu_i)^{\alpha^{-1}}$, and if ξ is a Logit, the resulting log-likelihood is given as,

$$\begin{aligned}\ln L(\theta,u,\beta,\gamma) = -\sum_{i=1}^{n} \ln\left(1+e^{q_i'u}\right) + \sum_{y=0} \ln\left(e^{q_i'u} + (1+\alpha\mu_i)^{\alpha^{-1}}\right) + \\ \sum_{y>0} \ln\left(\Gamma(y_i+\alpha^{-1})/\Gamma(y_i+1)\Gamma(\alpha^{-1})\right) - \\ (\alpha^{-1}+y_i)\ln(1+\alpha\mu_i) + y_i(\ln\mu_i + \ln\alpha)\end{aligned} \tag{A.8}$$

The mean of the ZI-NB2-Logit is given by $\nu_i = (1-\xi)\mu_i$ as before, and the variance is given by $\omega_i = (1-\xi)(1+\xi\mu_i+\theta\mu_i)\mu_i$.

Table A.1. MisProbit Vs ZI-MisProbit. Constant Misclassification

	MisProbit		ZI-MisProbit	
	Coefficient	Robust S.E	Coefficient	Robust S.E
Constant	-0.226	*(0.701)*	0.671	*(0.977)*
Immigrant	0.055	*(0.334)*	0.095	*(0.439)*
Age	-0.060	*(0.046)*	-0.126**	*(0.057)*
Age^2	0.000	*(0.000)*	0.001*	*(0.000)*
Male	0.926***	*(0.379)*	0.546***	*(0.212)*
White	0.826**	*(0.405)*	0.719*	*(0.388)*
Region South	0.147	*(0.250)*	0.484	*(0.493)*
Region Midlands	0.024	*(0.225)*	0.489	*(0.453)*
Region North	0.109	*(0.250)*	0.627	*(0.505)*
Prob of Misclassifying an One as Zero				
Constant	0.811***	*(0.066)*	0.480*	*(0.288)*
Prob of Misclassifying a Zero as One				
Constant	0.012***	*(0.005)*	0.014**	*(0.006)*
Prob of Zero Inflation				
Constant			2.509	*(6.468)*
Immigrant			0.136	*(0.977)*
Age			-0.075	*(0.053)*
Male			-0.448	*(0.383)*
White			-0.073	*(0.668)*
Region South			1.013	*(1.364)*
Region Midlands			1.283	*(1.179)*
Region North			1.359	*(1.216)*
Truthfulness			3.668	*(4.989)*
Sample Size	11,658		11,658	
Log Likelihood	-1,446.92		-1,429.16	
Predicted.Prob.Crime	0.280		0.269	
Predicted Prob. of Zero - Inflation			0.393	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level

Table A.2. MisProbit Vs ZI-MisProbit. Covariate-Dependent Misclassification

	(1)		(2)	
	Coefficient	Robust S.E	Coefficient	Robust S.E
Prob of Property Offence in Last Year				
Constant	2.335***	*(0.896)*	5.959	*(3.660)*
Immigrant	-0.254	*(0.273)*	0.756	*(0.748)*
Age	-0.259***	*(0.046)*	-0.386*	*(0.203)*
Age^2	0.003***	*(0.0006)*	0.003*	*(0.002)*
Male	0.440***	*(0.156)*	1.225	*(0.828)*
White	0.542*	*(0.293)*	1.380*	*(0.787)*
Region South	0.273	*(0.231)*	1.650	*(1.585)*
Region Midlands	0.124	*(0.228)*	1.973	*(1.902)*
Region North	0.496	*(0.318)*	1.951	*(1.642)*
Prob of Misclassifying an One as Zero (Underreporting)				
Constant	2.267***	*(0.507)*	-1.068**	*(0.505)*
Immigrant	-0.410	*(0.346)*	0.061	*(0.185)*
Age	-0.195***	*(0.074)*	0.165***	*(0.039)*
Age^2	0.003***	*(0.001)*	-0.003***	*(0.001)*
Male	-0.277*	*(0.142)*	-0.306***	*(0.085)*
White	-0.198	*(0.224)*	-0.394***	*(0.140)*
Region South	0.262	*(0.224)*	0.132	*(0.144)*
Region Midlands	0.224	*(0.215)*	0.305**	*(0.144)*
Region North	0.490**	*(0.212)*	0.208	*(0.155)*
Truthfulness	0.868***	*(0.256)*	0.623***	*(0.138)*
Prob of Zero Inflation				
Constant			-44.141	*(68.415)*
Immigrant			0.271	*(0.847)*
Age			8.640	*(12.355)*
Age^2			-0.406	*(0.533)*
Male			-0.557	*(0.498)*
White			0.354	*(0.718)*
Region South			-0.394	*(0.632)*
Region Midlands			-0.754	*(0.849)*
Region North			-0.503	*(0.948)*
Prob of Misclassifying a Zero as One				
Constant	-2.234***	*(0.204)*	-2.140***	*(0.083)*
Log Likelihood	-1,422.05		-1,415.86	
Predicted.Prob.Crime	0.292		0.509	
Predicted.Prob.Under	0.709		0.719	
Predicted.Prob.Inflation			0.076	

Robust standard errors are presented in parentheses.

(***), denotes statistical significance at 1% significance level

(**), denotes statistical significance at 5% significance level

(*), denotes statistical significance at 10% significance level