Immigration Status and Criminal Behavior

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

Georgios Papadopoulos*

School of Economics, University of East Anglia

February 2013

Abstract

This paper studies the individual level relationship between immigration and property crime in England and Wales using crime self-reports from the Crime and Justice Survey. Binary and count data models that account for under-reporting of criminal activity are used, since under-reporting is a major concern in self-reported crime data. The results indicate that under-reporting is considerably large, but, if anything, immigrants are less likely to under-report than natives. They also reveal that, once controlling for under-reporting and for basic demographic characteristics, even though not statistically significant, the effect of being an immigrant on crime is robustly negative across all model specifications (and statistically significant in some of those specifications). This might suggest that the negative association actually exists in the population, but the nature of the regression models in combination with the data in hand do not allow to estimate the relationship more precisely. We finally find that the effect of immigration status on property crime differs across regions and across ethnic groups.

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

*Address: School of Economics, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. E-mail: *G.Papadopoulos@uea.ac.uk*.
I am grateful to João Santos Silva, Tim Hatton, Zelda Brutti, Michail Veliziotis, Gianluigi Vernasca, Mariña Fernádez Salgado, Ken Burdett, and participants in the "Migration: Economic Change, Social Challenge" conference (UCL), RES 2010, SAEe 2010, EEA 2010, WPEG 2010, the "Economics of Culture, Institutions, and Crime" conference (FEEM), and the Crime Surveys User Meeting, ESDS, London, for their helpful comments. The empirical analysis is implemented in Stata/SE® and TSP® econometric software. The Crime and Justice Survey utilised in this project was commissioned by the Home Office and conducted by the BMRB Social Research and the National Centre for Social Research (NatCen). The 1995 British Social Attitudes survey (BSA) and the 2003,2009 BSA, also used in this project, were conducted by the Social and Community Planning Research and the NatCen, respectively. All data were provided by the UK Data Archive. Neither the original data collectors nor the aforementioned individuals bear any responsibility for the analyses or conclusions presented here.

1 Introduction

International migration is a topic that has been heavily debated by policy makers, especially in countries that experienced important immigration flows, such as the UK. As a result, academic communities have devoted extensive research to understand the actual impact of immigration on several aspects of both the host and home countries. These include the effect of immigration on the labour market (Borjas, 2003; Dustman, Fabbri and Preston, 2005; Card, 2009) and the welfare state of the host countries (Borjas, 1999), the impact of brain drain on the countries of origin (Beine, Docquier and Rapoport, 2008) and the impact of ethnic diversity on economic performance (Alesina and La Ferrara, 2005), to mention only a few. Following these substantial inflows and the debates, the general public started developing negative beliefs towards immigrant population, since they perceived immigrants not only as a major competitor in the labour market but also as a main factor for the deterioration of several problems of the host countries, especially crime. Indeed, at least for the UK, data from important social surveys, such as the British Social Attitudes survey (BSA), reveal that British citizens/nationals generally believe that immigrants increase crime rates (see, Appendix for details).

It comes as a surprise that although academics started debating the impact of immigration on crime more than 100 years ago (see, for example, Hart, 1896), only recently have researchers started investigating whether a relationship exists empirically. Very interestingly, a high proproportion of this recent research does not share the hostile view dominating individual attitudes. Actually, most investigations in the US indicate that if there is an association, this is negative (Bucher and Piehl, 1998, 2007; Lee, Martinez and Rosenfeld, 2001; Reid at al 2005; Ousey and Hubrin, 2009; Wadswarth, 2010), but evidence from Europe suggest mixed results for property crime but no association for violent crime (Bianchi, Buonanno and Pinotti 2008; Bell, Machin and Fasani, 2010; Bell and Machin, 2011).

Main objective of the present study is to shed more light on the differences in criminal behaviour between immigrants and natives in England and Wales, with a particular focus on property crime. For this purpose, the 2003 Crime and Justice Survey is used, a national representative survey of crime self-reports.[1] Under-reporting of criminal activities is a major concern in crime self-reports, even though the CJS uses techniques to improve the reliability of responses such as computer-based as opposed to face-to-face interviews (see for example, Turner at al 1998) and regression strategies that do not take into account this problem will result in inconsistent estimates of the determinants of criminal behaviour (see, Section 4). Therefore, in this study this problem is tackled by using regression models for both binary and count data that attempt to control for under-reporting. Given correct specification, these models consistently estimate both the determinants of true criminal activity and the determinants of reporting behaviour by using the information of observed self-reported crime only.

The results of these models suggest that responses of criminal behaviour are considerably

[1]For details on the survey design of the CJS refer to Hamlyn at al (2003). Note that Scotland and Northern Ireland are excluded from the CJS because of their separate criminal and justice system which generates incomparable crime statistics.

under-reported. Moreover, we find that under-reporting is not constant across individuals but a function of respondents' characteristics. However, if anything, immigrants under-report by less than natives. In addition, once controlling for under-reporting and for basic demographic characteristics, we find that on average, immigrants are not more or less involved in property crime, although the sign of the immigration status coefficient is robustly negative across all model specifications. Furthermore, although we focus on property crime, violent crime results are also presented suggesting a negative, but statistically insignificant, association between immigration status and crime as well. All the robustness checks show that although the estimated immigration-crime difference is not statistically significant, it is actually quite robust, which may suggest that this relationship exists, but the nature of the regression models in combination with the data in hand do not allow for more precise estimates.

Next, we investigate whether the effect of immigration status on crime depends on certain groups of covariates, such as ethnic status or location. For example, there is high heterogeneity among immigrants in terms of ethnicity and therefore we might expect different ethnic groups, perhaps due to cultural factors, to differ in their criminal activity. Furthermore, location of immigrants is not randomly assigned, but it is a rather complicated process that depends on many factors that may be related to criminal activity. For example, if immigrants try to match their abilities with the opportunities that each area provides, more crime-prone immigrants will decide to locate in areas that offer more criminal opportunities. The results actually suggest that the effect of immigration on property crime depends on their ethnic status (Black immigrants being the least criminally involved ethnic group) and their location (London being the area with the most law-abiding immigrants).

The remainder of this paper is organised as follows. Section 2 puts the individual decision to commit property crimes in a simple economic framework of individual supply of crime. Utilising this simple economic model, it also investigates the individual relationship between immigration and property crime. Section 3 presents a brief review on the immigration-crime literature. Section 4 offers a presentation of the regression models that intend to control for potential under-reporting. Section 5 discusses the data and the variables. It also offers some basic descriptive statistics. The main results follow in Section 6 with robustness checks following in Section 7. Section 8 investigates whether the effect of immigration status on property crime depends on ethnic status or the location of immigrants. Finally, a brief discussion and concluding remarks follow in Section 9.

2 An Economic Model of Participation in Property Crime Activities

According to economic theory there are generally two channels through which immigration can have an impact on crime. The first one, namely the "indirect effect", states that flows of immigration affect crime rates due to their influence on labour market outcomes of the domestic economy which in turn are related to criminal activities. However, at least for the Great Britain, it has been generally found that there are minimal effects of immigration on the British labour market outcomes (see for example, Dustman, Fabbri and Preston, 2005; Manacorda, Manning and Wadsworth, 2012).[2] The second one, namely the "direct effect", states that immigrants might be more or less crime-prone than natives, since there are differences in their characteristics associated with criminal activities, such as differences in labour opportunities or risk attitudes. As this paper focuses on the individual level relationship between immigration and crime, this section investigates only the latter by developing a simple economic model of crime supply. Although this model focuses on property crime, it is also applicable to violent crime with slight modifications.

This is a one period model under uncertainty that borrows features from Ehrlich (1973), and Lochner and Moretti (2001). Although this is not a complete investigation of criminal behaviour, it well illustrates why differences in participation in illegitimate activities between immigrants and natives may exist. Consider a rational individual who, conditionally on time spent in leisure, optimally decides how to allocate his/her available time, τ, between legal and illegal activities, denoted as τ_ℓ and τ_i respectively. Immigration status is given by the binary indicator m, so that $m = 1$ if the individual is an immigrant and $m = 0$ otherwise.

If the individual participates in the legal sector, he/she can be either employed (State A) or unemployed (State B) depending on the, exogenously given, probability of unemployment $\mu(m)$. If employed, he/she receives wage $w(\tau_\ell, m)$ with $dw(\cdot)/d\tau_\ell > 0$, whereas if unemployed, he/she receives the unemployment benefit $D(\tau_\ell)$ with $dD(\cdot)/d\tau_\ell > 0$. It is also assumed that $\underline{w}(\tau_\ell) > D(\tau_\ell)$ and $d\underline{w}(\cdot)/d\tau_\ell > dD(\cdot)/d\tau_\ell$, where $\underline{w}(\tau_\ell)$ is the minimum wage rate. On the other hand, if the individual participates in the illegal sector, he/she receives the "criminal wage" $k(\tau_i, m)$, which consists of financial and psychological outcomes measured in their monetary equivalent, with $dk(\cdot)/d\tau_i > 0$. Thus, psychological costs associated with crime, such as bad reputation, regret, uneasiness, etc, are incorporated in $k(\tau_i, m)$. We assume that illegal opportunities that pay high pecuniary returns require considerable time in the illegal sector or/and they involve higher psychological costs. In addition, if the individual spends time on committing crime, he/she also face the probability of apprehension, $\pi(\tau_i, m)$, with $d\pi(\cdot)/d\tau_i > 0$, and if apprehended, punishment, $P(\tau_i, m)$, occurs with certainty (without loss of generality), with $dP(\cdot)/d\tau_i > 0$. Punishment is also measured in its monetary equivalent and happens at the end of the period, so that the individual discounts it by a discount rate $\rho(m)$.

If we assume for simplicity that expected punishment is measured in utility terms as in Lochner and Moretti (2001), the expected utility gained from both legal and illegal activities is given by,

$$U(\tau_i, \tau_\ell) = (1 - \mu(m))\, u\,(y_a) + \mu(m)\, u\,(y_b) - \rho\, \pi(\tau_i, m) P(\tau_i, m), \tag{1}$$

[2] Note that criminologists have also developed several theories, which suggest that immigration might influence crime rates as it may impose cultural conflicts and cause social disorganisation (see, for example, Martinez and Lee, 2000).

where, $y_a = w(\tau_\ell, m) + k(\tau_i, m)$ and $y_b = D(\tau_\ell) + k(\tau_i, m)$, are the returns from State A and State B respectively. Finally, assume that $u'(y_\jmath) > 0$, and $u''(y_\jmath) < 0$, where $\jmath = (a, b)$. Henceforth, m is omitted from the equations for brevity. Thus, the problem of the individual is to allocate his/her available time between legal and illegal activities in order to maximize (1) subject to the time constraints, $\tau = \tau_i + \tau_\ell$, and, $\tau_i \geq 0, \tau_\ell \geq 0$. The Lagrangian function is therefore given by,

$$\mathcal{L}(\tau_i, \tau_\ell, \lambda) = (1 - \mu(m))\, u\,(y_a) + \mu(m)\, u\,(y_b) - \rho\, \pi(\tau_i, m) P(\tau_i, m) + \lambda(\tau - \tau_i + \tau_\ell), \tag{2}$$

where λ is the Lagrange multiplier, and the Kunh-Tucker first order conditions are,

$$\frac{dU\,(\tau_i)}{d\tau_i}\tau_i = 0, \quad \frac{dU\,(\tau_i)}{d\tau_i} \leq 0, \quad \tau_i \geq 0; \qquad \frac{dU\,(\tau_\ell)}{d\tau_\ell}\tau_i = 0, \quad \frac{dU\,(\tau_\ell)}{d\tau_\ell} \leq 0, \quad \tau_\ell \geq 0. \tag{3}$$

The interior solution is obtained when $dU(\tau_i^*)/d\tau_i^* = 0$ and $dU(\tau_\ell^*)/d\tau_\ell^* = 0$, which can be expressed as,

$$\Big((1 - \mu)u'(y_a) + \mu\, u'(y_b)\Big)\frac{dk(\cdot)}{d\tau_i^*} - \left((1 - \mu)u'(y_a)\frac{dw(\cdot)}{d\tau_\ell^*} + \mu\ u'(y_b)\frac{dD(\cdot)}{d\tau_\ell^*}\right)$$
$$= \rho\left(\frac{d\pi(\cdot)}{d\tau_i^*}P(\cdot) + \frac{dP(\cdot)}{d\tau_i^*}\pi(\cdot)\right), \tag{4}$$

so that the marginal utility obtained from criminal activities minus the marginal utility obtained from legal activities must be equal to the marginal punishment.[3] Since the RHS of (4) is weakly positive, the individual will spend some time on illegal activities if and only if the marginal utility from criminal activities is at least as high as the marginal utility from the legal sector. Therefore, the RHS of (4) can be considered as the marginal compensation required to cover for the risk of spending time on committing crimes.

As the criminal wage rate is in general quite small in comparison to the legal wage rate for most property crimes, and if we consider that for most people the criminal wage further decreases by the psychological costs associated with a crime, the corner solution where someone allocates all his/her time in legal actions is highly possible. Property crimes that pay a high financial return are also very rare, as they require plenty of time which in turn increases the risk of apprehension and the severity of punishment, or because they involve much higher psychological costs than psychological gains for most people. Finally, most individuals do not exhibit strong criminal ability which might decrease $k(\tau_i)$ (if less able criminals target in criminal activities that pay low returns) or increase $\pi(\tau_i)$. On the other hand, the individual will specialise in the illegal sector, if and only if the marginal legal utility plus the marginal cost of punishment is smaller than the marginal utility from illegitimate activities, which is highly unlikely for most people.

What could (4) and a simple comparative statics analysis tell us about differences in the criminal activities between the average immigrant and the average native? Since m appears in most determinants of (4), immigration status affects criminal behaviour through many channels.

[3]For the sufficient condition for a strict global maximum refer to Papadopoulos (2011).

First of all, it is clear from (4) that, starting from an equilibrium where the individual spends some time on the illegal sector, an increase in the marginal utility gained from the legal sector will decrease the LHS of (4) and therefore, holding everything else constant, the participation in illegal activities becomes less likely (it increases the opportunity cost of crime), and vice versa for an increase in the marginal utility gained from the illegal sector. In addition, the effect of an increase in unemployment rate, μ, increases participation in crime as somebody would expect. The comparative static analysis shows that $d\tau_i^*/d\mu > 0$ if and only if, $\frac{u'(y_b)}{u'(y_a)} > \frac{(dk/d\tau_i - dw/d\tau_\ell)}{(dk/d\tau_i - dD/d\tau_\ell)}$. Now, as $\frac{dD}{d\tau_\ell} < \frac{dw}{d\tau_\ell}$, the RHS of this inequality is lower than one. Moreover, since $w(\tau_\ell) > D(\tau_\ell)$, then $y_a > y_b$, and therefore $u'(y_b) > u'(y_a)$ due to strictly concavity of $u(\cdot)$. Thus, the LHS of this inequality is higher than one and consequently this inequality always holds. Since immigrants on average face lower legal opportunities, such as lower $dw(\cdot)/d\tau_\ell$, or higher μ (see, for example, Algan at al, 2010), we would expect immigrants to be more crime-prone than natives. Regarding criminal opportunities, there is no evidence on whether or not immigrants face a higher or a lower $dk(\cdot)/d\tau_i$ than natives and therefore it would not be appropriate to associate immigration with crime using the return to criminal activities.

In addition, an exogenous increase in $\pi(\tau_i)$ or $P(\tau_i)$ decreases participation in crime as expected, since it increases the RHS of (4). At the same time we would expect that the typicall immigrant faces higher $\pi(\tau_i)$ and $P(\tau_i)$ than the typical native. Firstly, 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. Secondly, there is some evidence suggesting that the criminal justice system is biased in various stages against ethnic minorities (Smith, 1997, Feilzer and Hood, 2004). This implies that, for the same crimes, immigrants may face more severe punishments compared to natives. Lastly, immigrants also face deportation which is a punishment specific to them and can be considered as an important disincentive to commit crimes (Butcher and Piehl, 2007). Thus, according to the above, a negative association between immigration status and criminal behaviour can be expected.

Finally, risk attitudes, which can be expressed through the discount factor or the curvature of the utility curves, are quite important on determining criminal behaviour. For example, people that are very "patient" discount future potential punishment less heavily (higher ρ) which increases the RHS of (4). Moreover, more risk averse individuals are represented by a more "curvy" utility functions. Thus, as y goes up, $u'(.)$ decreases by more for a more risk averse individual, which, ceteris paribus, results in a smaller marginal utility gained from both legal and illegal activities (LHS of (4) becomes smaller). In both cases, a higher marginal compensation is required to cover for the extra risk. According to this, we would expect differences between immigrants' and natives' participation in criminal activities because discount factors and risk attitudes may as well be quite different between them. For instance, since migration involves a lot of uncertainty, it could be said that people who decide to emigrate are willing to take higher risks than those who decide not to (see, for example, Jaeger at al, 2010). However, this is not informative at all with respect to the risks they are willing to take compared to natives once they arrive at the host country. For example, empirical evidence suggests that immigrants are actually more risk averse than natives (see, for example, Bonin at al, 2009).

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. However, as will be discussed in Section 6, the covariates we use in the regression analysis control for these differences in basic demographics.

Taking all the above discussion into consideration, simple economic theory predicts that the effect of immigration status on criminal activity can go either way, and therefore it should be established by an empirical investigation using a well specified model and appropriate data.

3 A Brief Literature Review on the Relationship between Immigration and Crime

As mentioned in the introduction, research from the US indicates that in general immigration has a negative impact on crime, but evidence from Europe suggests mixed results for property crime but no association for violent crime. Although there is some agreement on the statement above, the empirical evidence is inconclusive, as different studies report very different results, depending on the host countries being studied, the composition of immigrants and the circumstances immigrants encounter on those countries, the differences in data sources, but most importantly, the statistical tools and strategies each researcher uses to answer the research question.

Most researchers have used administrative crime panel data on an attempt to estimate the aggregate (macro-)causal impact of immigration on crime by relating changes in migration stocks to changes in crime rates. Major problem in identifying a causal impact, as opposed to merely associations, is that location of immigrants is endogenous. For example, immigrants are disproportionally located in deprived areas where crime is higher, just because they cannot afford expensive housing or because they tend to locate where there is a large population of residents of the same ethnic background. Although panel data techniques, such as the fixed effect estimator (see, Bucher and Piehl, 1998; Ousey and Hubrin, 2009; Wadswarth, 2010), alleviate this problem by controlling for unobserved time-invariant location characteristics, it does not solve it completely, as there might always be unobserved time-varying features that affect the decision of location and crime rates at the same time. Some researchers tried to deal with this problem by using instrumental variables techniques, trying to find exogenous variations that affect immigrants' choice of location but not unobserved factors that affect crime rates (see, Bianchi, Buonanno and Pinotti 2008; Bell, Machin and Fasani, 2010; Spenkuch, 2010).

Of particular interest are the papers by Bell, Machin and Fasani (2010) and by Bell and Machin (2011), as to my best knowledge these are the only papers that use data from the UK. Bell, Machin and Fasani (2010) examine the impact on crime of two separate large waves of immigrants, the late 1990s wave of asylum seekers and the large inflow from the "A8" Eastern European countries since May 2004. They find that the first wave is associated with higher property crime, even after controlling for endogenous location using fixed effects and instrumental variables.[4] However, they find that the A8 wave did not affect property crime.[5] 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. Bell and Machin (2011) study the effect of immigrant neighbourhood segregation on crime using both police recorded data and self-reports from the British Crime Survey (BCS). They find strong evidence across different specifications that, ceteris paribus, crime is lower in enclaves, where an enclave is defined as a neighbourhood with at least 30% immigrant population. In both papers, the findings suggest that there is no effect for violent crime.

On another direction, researchers focused on the direct (micro-)impact by studying whether immigrants are more or less crime-prone than natives. Official criminal records, such as arrests, convictions, or prison records, have been a popular choice among researchers (see, Albrecht, 1997; Kilias, 1997; Martens, 1997; Yeager, 1997). According to this type of data, immigrants are usually overrepresented on official statistics, meaning that the proportion of immigrants in the imprisoned/convicted population is higher than the proportion of immigrants in the general population. The main flaw with this strategy is that official records provide a distorted picture of actual crime. In fact, there is strong empirical evidence that a high proportion of committed crimes remain unrecorded by the police, - the so-called "dark figure" of crime (see, MacDonald, 2002) - either because victims for several reasons do not report them or due to reasons that have to do with police recording practices. If the recording mechanism is somehow biased against immigrants, then these statistics actually overestimate immigrants' involvement into criminal activities compared to natives. For example, it has been argued 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), so that immigrants might face a higher probability of arrest or conviction for the same crimes relative to natives, or, police officers disproportionately record crimes that have been supposedly committed by immigrants. Immigrants may also be more "visible" to the police because of over-policing in target areas where ethnic minorities are concentrated, increasing the likelihood of immigrants to be arrested relative to natives (Sharp and Budd, 2005).

Another common practice among researchers has been the use of crime self-reports.[6] It is interesting that, as opposed to official records, researchers generally find that immigrants are more law abiding than natives (see, for example, Junger-Tas, 1997; Bucher and Piehl, 1998; Vazsonyi and Kilias, 2001). Although relying on self-reported crime has been the most popular strategy used to determine causes of criminal or antisocial behaviour (see, Junger-Tas & Marshall

[4]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.

[5]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.

[6]For details on the methodology of crime self-report studies see, Junger-Tas and Marshall (1999); Thornberry and Krohn (2000).

1999), reliability of responses is a major concern, since questions try to elicit information on a very sensitive part of personal behaviour. Therefore, a response error is expected that most probably takes the form of under-reporting. Thus, if for some reasons immigrants tend to under-report by more than natives, the estimated differences reported by the studies above reflect to some extent differences in reporting behaviour rather than differences in criminal activity. To the best of my knowledge, no study using crime self-reports has attempted to somehow correct for possible misreporting.

4 Regression Models

Before starting our analysis, it is important to point out that, as we will see in Section 5, the observed information on property crimes is given in count form (y_p = number of property crimes committed during the twelve months prior to the interview). Therefore, nonlinear regression models for count data are more appropriate than linear ones. Conventional nonlinear estimators are inconsistent if under-reporting, or more generally, response error in the outcome variable is present (see, Hausman, Abrevaya and Scott-Morton, 1998; Cameron and Trivedi, 1998; Winkelmann, 2008). The problem is even more salient if under-reporting depends on individual characteristics which is certainly the case with crime self-reports. Therefore, models that take into account misreporting should be used.

However, from the probability distribution of y_p presented in Table 1 (see, Section 5), and its unconditional (weighted) mean and standard deviation presented in Table 3 (see, Section 5), we notice two important features that complicate estimation of count data models. Firstly, there is an exceptionally large concentration on outcome zero since 94.17% of the respondents reported that they committed zero crimes. Secondly, positive outcomes are highly dispersed with some extreme values. Due to these two features, estimation of count data models that allow for under-reporting is quite unstable. Alternatively, a "safer" choice is to utilise estimators that use only the binary choice information, whether or not someone has committed a crime last year, even though some information is neglected, and study whether immigrants are more or less likely than natives to commit property crimes once controlling for misreporting and basic demographics. Therefore, the main results of this research work are based on a parametric binary choice estimator that takes into account both under-reporting (misclassification of a true one as zero) and over-reporting (misclassification of a true zero as one), presented in subsection 4.1, while parametric count data models that allow for under-reporting, presented in subsection 4.2, are used for sensitivity analysis.

4.1 The Misclassification Probit Model

The Misclassification Probit model presented here, henceforth called the MisProbit, is developed by Hausman, Abrevaya and Scott-Morton (1998). It arises naturally from a latent variable specification. Assume that in a given period of time, an individual will spend some time on committing crimes if the net utility from committing these crimes is positive. So, let U_i^* be the (unobserved) utility obtained if committing these crimes minus the utility if not committing them and assume that U_i^* is linear function of the $(K \times 1)$ vector of covariates $\mathbf{x}_i$ such that,

$$U_i^* = \mathbf{x}_i'\boldsymbol{\beta} + \epsilon_i, \quad \text{for, } i = 1, \dots n, \quad \text{and, } \epsilon_i|\mathbf{x}_i \sim \mathcal{N}(0,1). \tag{5}$$

Thus, the individual commits at least one crime according to,

$$y_i^* = \begin{cases} 1 & \text{if } U_i^* > 0 \\ 0 & \text{if } U_i^* \leq 0, \end{cases} \tag{6}$$

where y^* denotes the binary variable for the true but unobserved crime. Therefore, conditional on $\mathbf{x}_i$, the probability of committing a crime is given by,

$$\Pr(y_i^* = 1|\mathbf{x}_i) = \Pr(U_i^* > 0|\mathbf{x}_i) = \Pr(\epsilon_i > -\mathbf{x}_i'\boldsymbol{\beta}|\mathbf{x}_i) = \Phi(\mathbf{x}_i'\boldsymbol{\beta}) \tag{7}$$

where $\Phi(\mathbf{x}_i'\boldsymbol{\beta})$ is the standard normal CDF.

Suppose now that the reported and therefore observed crime, denoted by the binary variable y_i, does not coincide with y_i^* since there is misreporting, in this context defined as misclassification. The probabilities of misclassification are defined as follows,

$$a_0 = \Pr(y_i = 0|y_i^* = 1), \quad \text{and,} \quad a_1 = \Pr(y_i = 1|y_i^* = 0), \tag{8}$$

where a_0 is the probability of reporting zero crimes conditional on committing at least one crime (under-reporting) and a_1 is the probability of reporting a crime, conditional on committing zero crimes (over-reporting). Notice that according to this specification the misclassification probabilities do not depend on $\mathbf{x}_i$ but only on y_i^*. Following the response tree in Figure 1(a), the conditional probability of observing a crime is given by,

$$\Pr(y_i = 1|\mathbf{x}_i) = (1 - \Phi(\mathbf{x}_i'\boldsymbol{\beta}))\, a_1 + \Phi(\mathbf{x}_i'\boldsymbol{\beta})(1 - a_0) \;=\; a_1 + (1 - a_0 - a_1)\Phi(\mathbf{x}_i'\boldsymbol{\beta}), \tag{9}$$

which is also the conditional expectation.

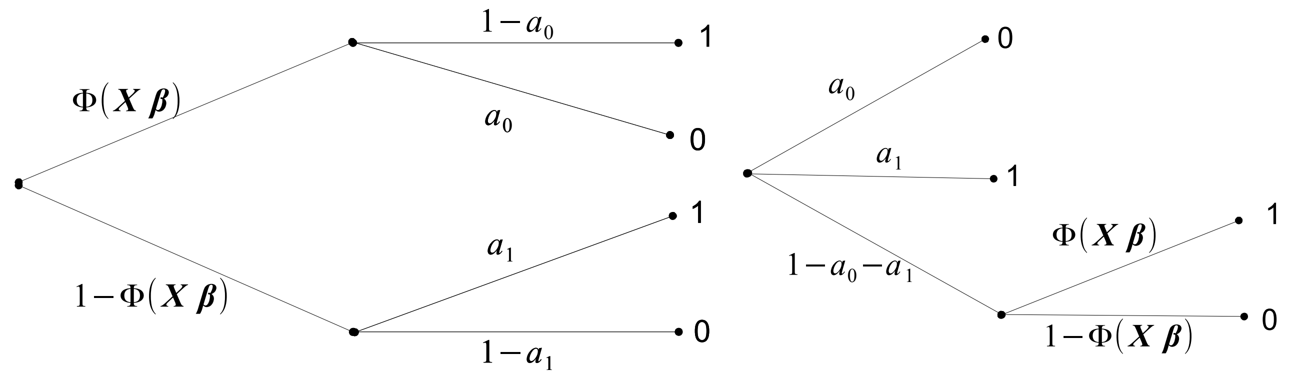
We can estimate, $\boldsymbol{\theta}_B = (a_0, a_1, \boldsymbol{\beta})$, (where B stands for 'binary model') using the Maximum Likelihood Estimator (MLE) since the log-likelihood function is simply given by,

$$\ln \mathcal{L}(\boldsymbol{\theta}_B) = \sum_{i=1}^{n} \Big(y_i \ln\left[\Pr(y_i = 1|\mathbf{x}_i)\right] + (1 - y_i) \ln\left[1 - \Pr(y_i = 1|\mathbf{x}_i)\right] \Big). \tag{10}$$

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 (10) using optimization algorithms, such as the Newton-Raphson, yields the vector of consistent estimates $\hat{\boldsymbol{\theta}}_B$, where $\hat{\boldsymbol{\beta}}$ are the estimated coefficients of true crime, $\hat{a}_0$ is the estimated probability of under-reporting, and $\hat{a}_1$ is the estimated probability of over-reporting.

Figure 1: **Probability Trees**

(a) Misclassification (b) Zero-One Inflation



There are several issues we need to stress at this point. First of all, Hausman, Abrevaya and Scott-Morton (1998) show that $\boldsymbol{\theta}_B$ is not identified since for a symmetric CDF, $F(\cdot)$, $a_1 + (1 - a_0 - a_1)F(\mathbf{x}_i'\boldsymbol{\beta}) = \tilde{a}_1 + (1 - \tilde{a}_0 - \tilde{a}_1)F(-\mathbf{x}_i'\boldsymbol{\beta})$ where $\tilde{a}_0 = 1 - a_1$ and $\tilde{a}_1 = 1 - a_0$, which means that there are two observationally equivalent models with parameters $(a_0, a_1, \boldsymbol{\beta})$ and $(\tilde{a}_0, \tilde{a}_1, -\boldsymbol{\beta})$. 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 $\tilde{a}_0 + \tilde{a}_1 > 1$. If this condition fails, the misclassification probabilities are too large, and therefore, the data are most probably too noisy to obtain reasonable results.

Note also that a_0 is designed to capture only "total" as opposed to "partial" under-reporting. That is, the probability of under-reporting will ignore the cases where individuals report just a proportion of the total number of crimes they have committed.

In addition, we need to stress that exactly the same model can be obtained under a zero-one-inflation framework. That is, a proportion of individuals, a_0, regardless of $\mathbf{x}_i$, never commits any crimes and another proportion a_1, regardless of $\mathbf{x}_i$, always commits at least one crime. According to this framework, if we assume that there is no misreporting, it is clear from Figure 1(b) that the probability of observing a crime is given as in (9). As a result, the model cannot distinguish under-reporting from zero-inflation and over-reporting from one-inflation. Although probability of one-inflation is unrealistic, zero-inflation is possible and therefore, a_0 must be interpreted with caution. Although it is possible to separate under-reporting from zero-inflation by incorporating zero-inflation separately in the likelihood function, this model is not discussed further here (for a discussion of this model together with some results refer to Papadopoulos, 2011).

Most importantly, the assumption of constant misclassification is not realistic in our application. Nevertheless, this assumption can be easily relaxed if we model a_0 and a_1 to be functions of covariates, so that, $a_{0i} = F(\mathbf{z}_{0i}'\boldsymbol{\gamma}_0)$ and $a_{1i} = F(\mathbf{z}_{1i}'\boldsymbol{\gamma}_1)$, where F can be the CDF of a binary model, such as a Probit or a Logit. Note that $\mathbf{x}, \mathbf{z}_1$ and $\mathbf{z}_2$ may be identical, overlapping or disjoint, but what variables are actually included in these vectors depends on the particular application under investigation.

Although this model is easy to implement in statistical software, in practice estimation can be difficult, particularly when the misclassification probabilities are allowed to depend on variables, when $\mathbf{x}, \mathbf{z}_1$ and $\mathbf{z}_2$ contain the same variables, and when the proportion of zero outcomes is very high (remember that around 94% of respondents reported zero crimes). Exclusion restrictions could always facilitate the estimation procedure. Moreover, as Hausman, Abrevaya and Scott-Morton (1998) note, the MisProbit is less efficient than the usual Probit model since its Information Matrix (given by $-E\left(\partial \ln \mathscr{L}(\boldsymbol{\theta}_B)/\partial \boldsymbol{\theta}_B \partial \boldsymbol{\theta}'_B\right)$) is not block-diagonal, even in the absence of misclassification. Thus, the MisProbit produces higher standard errors, while this loss in precision increases as the misclassification probabilities increase. However, notice that if the MisProbit is the correct specification, the estimated standard errors correspond to the true imprecision while the conventional Probit actually produces underestimated standard errors. According to the above, in cases where misclassification is very high, as it may be the case for crime self-reports, quite rich samples are needed in order to obtain precise estimates.

4.2 The NB2-Logit and the ZI-NB2-Logit

Here we briefly present and discuss the Negative-Binomial(2)-Logit (NB2-Logit) model developed in Winkelmann and Zimmermann (1993) and a generalisation of it that incorporates zero-inflation. For a more detailed analysis see Papadopoulos (2011). We prefer the NB2 distribution to the usual choice of the Poisson, since the equidispersion assumption ($E(y^*|\mathbf{x}_i) = Var(y^*|\mathbf{x}_i)$) imposed by the Poisson is too restrictive for our application. As an indicator, we can look at the unconditional (weighted) sample mean and variance of the count of observed crime which are 0.34 and 34.32 respectively, even though we are interested in their conditional counterparts, .

To begin with, for individual i (with $i = 1, \ldots, n$), suppose that y_i^* measures the number of committed crimes. We assume that y^* conditional on a vector of covariates $\mathbf{x}_i$ follows the Negative-Binomial(2) (NB2) with, $E(y_i^*|\mathbf{x}_i) = \lambda_i = \exp(\mathbf{x}_i'\boldsymbol{\delta})$ and $Var(y_i^*|\mathbf{x}_i) = \omega_i = \lambda_i + \alpha\lambda_i^2$, where α captures gamma specific unobserved heterogeneity, with $\alpha > 0$ and therefore $\omega_i > \lambda_i$.[7] However, only a subset of y_i^* is actually observed since i might decide not to report a number of committed crimes. Estimating a conventional regression model for count data based on the observed crime only, will ignore under-reporting and consequently will result in inconsistent estimates for the determinants of true crime. In this model, the total number of observed crimes, y_i, is given by the sum of a sequence of IID Bernoulli variables, c_{ij} ($j = 1, \ldots, y^*$) with a constant probability of 'success' p_i, where c_{ij} denotes a particular crime j committed by i. Thus, we can write,

$$y_i = c_{i1} + c_{i2} + \ldots\ldots + c_{iy^*} = \sum_{j=1}^{y_i^*} c_{ij}. \tag{11}$$

In our regression framework the probability to report a committed crime, p_i, follows the Logit model with $p_i = \Pr(c_{ij} = 1|\mathbf{z}_i) = \Lambda(\mathbf{z}_i'\boldsymbol{\eta}) = \exp(\mathbf{z}_i'\boldsymbol{\eta})/(1 + \exp(\mathbf{z}_i'\boldsymbol{\eta}))$.

[7] For details on the NB2 and other distributions for count or other discrete variables distributions refer to Johnson, Kemp, Kotz (2005). For count data regression models refer to Cameron and Trivedi (1998) or Winkelmann (2008).

If we further assume that y_i^* and c_{ij} are conditionally independent, it is easy to show, using probability generating functions as in Feller (1968), that the distribution of y_i conditional on $\mathbf{x}_i$ and $\mathbf{z}_i$ is also NB2 with modified mean and modified variance equal to,

$$\mu_i = \lambda_i p_i = \exp(\mathbf{x}_i'\boldsymbol{\delta})\frac{\exp(\mathbf{z}_i'\boldsymbol{\eta})}{1+\exp(\mathbf{z}_i'\boldsymbol{\eta})} \quad \text{and,} \quad \omega_i = \mu_i + \alpha\mu_i^2, \tag{12}$$

respectively. This is the NB2-Logit model, with PDF given by,

$$\Pr(Y_i = y_i|\mathbf{x}_i, \mathbf{z}_i) = \frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}(1+\alpha\mu_i)^{-(\alpha^{-1}+y_i)}(\alpha\mu_i)^{y_i}. \tag{13}$$

We can estimate $\boldsymbol{\theta}_C = (\boldsymbol{\delta}, \boldsymbol{\eta}, \alpha)$ (where C stands for "count model") using the MLE since we can easily obtain the the log-likelihood function as,

$$\ln \mathscr{L}(\boldsymbol{\theta}_C) = \sum_{i=1}^{n}\left(\ln\left(\frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}\right) - (\alpha^{-1}+y_i)\ln(1+\alpha\mu_i) + y_i(\ln\mu_i + \ln\alpha)\right). \tag{14}$$

Maximization of (14) yields consistent estimates of $\boldsymbol{\theta}_C$, given correct specification of the model; that is, the true DGP is NB2-Logit.[8] Note that, again, depending on the application under investigation, $\mathbf{x}_i$ and $\mathbf{z}$ may be identical, overlapping or disjoint. However, as Papadopoulos and Santos Silva (2008) show, unless appropriate restrictions are imposed, identification of $\boldsymbol{\theta}_C$ is not possible since there are two different sets of parameters leading to the same likelihood value. As suggested by Papadopoulos and Santos Silva (2008), one way to identify the parameters of interest is to impose at least one "strong" exclusion restriction on the NB2 part (crime process), meaning that there is at least one variable that has a signficant impact on the reporting process but no impact on the crime process. If such a variable exists, even though we still have (at least) two local maxima, the "correct" set of estimated parameters always leads to the highest maximum. If the effect of this variable is statistically very close to zero, the two sets of estimates lead to likelihood values that are approximately the same, and identification is questionable.[9] Finally, note also that the NB2-Logit model assumes that there is no over-reporting of crime.

This model can be easily extended to take into account zero-inflation giving rise to the ZI-NB2-Logit model. Now, let ξ be the probability of being an individual who never commits and consequently never reports crimes. Therefore, there is probability $(1-\xi)$ to be an individual who may commit crimes, but his/her responses are also subject to under-reporting, meaning that they follow the NB2-Logit model. In our regression framework the probability of zero-inflation, conditional on a set of characteristics $\mathbf{q_i}$, also follows a Logit model with $\xi_i = \exp(\mathbf{q}_i'\boldsymbol{\vartheta})/(1 + \exp(\mathbf{q}_i'\boldsymbol{\vartheta}))$.

[8]Correct specification of the mean only is not enough for consistency of $\hat{\boldsymbol{\theta}}_C$, as the NB2-Logit belongs to the Linear Exponential Family (LEF) only for a given value of α (see, Gourieroux, Monfort, and Trognon, 1984). However, since α is subject to estimation, NB2-Logit is not an LEF and therefore, misspecification of higher moments than the mean leads to inconsistency.

[9]Note also that, alternatively, the model can be identified imposing a sign restriction on the reporting process, meaning that we know with certainty the sign of one element of $\boldsymbol{\eta}$. However, since in the current empirical study this information is not available, this possibility is not discussed further.

According to this model, the conditional probabilities of zero and positives outcomes are given by,

$$\begin{aligned} \Pr(y_i = 0|\mathbf{x}_i, \mathbf{z}_i, \mathbf{q}_i) &= \xi_i + (1-\xi_i)(1+\alpha\mu_i)^{\alpha^{-1}}, \\ \Pr(y_i > 0|\mathbf{x}_i, \mathbf{z}_i, \mathbf{q}_i) &= (1-\xi_i)\frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}(1+\alpha\mu_i)^{-(\alpha^{-1}+y_i)}(\alpha\mu_i)^{y_i}, \end{aligned} \tag{15}$$

respectively, where $(1+\alpha\mu_i)^{\alpha^{-1}}$ is the probability of a zero outcome from the NB2-Logit model and $\mu_i = \lambda_i p_i$. Note that ξ_i, as was the case for a_0 in the binary choice model, cannot distinguish between the zero outcome because of zero-inflation or due to total under-reporting. The log-likelihood function is given by,

$$\begin{aligned} \ln \mathscr{L}(\boldsymbol{\theta}_C, \boldsymbol{\vartheta}) = -\sum_{i=1}^{n} \ln\left(1+\exp(\mathbf{q}_i'\boldsymbol{\vartheta})\right) + \sum_{y=0} \ln\left(\exp(\mathbf{q}_i'\boldsymbol{\vartheta}) + (1+\alpha\mu_i)^{\alpha^{-1}}\right) + \\ \sum_{y>0} \ln\left(\frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}\right) - (\alpha^{-1}+y_i)\ln(1+\alpha\mu_i) + y_i(\ln\mu_i + \ln\alpha). \end{aligned} \tag{16}$$

Identification of this model requires the same assumptions established for the NB2-Logit and consistency requires that the true DGP is ZI-NB2-Logit.

5 The CJS Data and Discussion of Variables

To investigate whether immigrants are more or less crime-prone than natives, the Crime and Justice Survey (CJS) of 2003 is used, a national representative study where respondents from England and Wales were asked questions about their criminal activities with respect to both property and violent crimes. Even though computer-based interviews were used, a method proven to increase reliability of responses (see, Turner at al, 1998), respondents may have still been reluctant to reveal information about their criminal activities. Therefore, some degree of under-reporting in the data is expected. As a response to this, this paper uses the regression models discussed in the previous section. To have a rough idea about the level of under-reporting in the CJS, we could compare the crime figures obtained from the CJS with crime figures from the British Crime Survey (BCS), since it is generally agreed that the BCS provides a relatively precise picture of criminal activity. For example, Budd, Sharp and Mayhew (2005) suggest that the figure of violent crime from the CJS is quite close to that of BCS, but the count of property crime is quite lower than in BCS. However, they also point out that these figures must be treated with caution since there are fundamental design differences between these two surveys.

Even though response rates of the CJS are very close to response rates of other surveys,[10] such as the Labour Force Survey or the BCS (see, Sharp and Budd, 2005), there is also a potential sample selection problem, since it is likely that people who refused participation in CJS, or

[10] The response rate for the core sample and youth-boost sample is around 74%. For the nonwhite-boost sample the response rate is around 50% which is common in surveys that include nonwhite boosts.

people who were incarcerated when the CJS was taking place, are likely to be more prone to crime than participants. However, models that correct for sample selection problems, such as the Heckit procedure (see, Heckman 1979), would require information on non-respondents, which is not available. Therefore, this problem is ignored in the analysis, hoping that the selection is on "covariates" as defined in Wooldridge (2007).

The information on property crime used to construct the dummy PROPERTY_B (taking value 1 if the respondent committed a property crime during the twelve months prior to the interview and 0 otherwise) and the count variable PROPERTY_C (number of property crimes committed during the same reference period) consists of: vehicle related thefts, domestic and commercial robbery, domestic and commercial burglary, thefts from person, thefts from work, school, shops, other thefts, and criminal damage of cars or other objects. The distribution of the property crime count variable is presented in Table 1, while descriptive statistics are provided in Table 3.

- TABLE 1 ABOUT HERE -

It is worth noting at this point that some effort is made to keep the sample size as large as possible. This is because, as explained in the previous section, not only are the econometric methods used in the empirical analysis quite demanding, but also around 94% of respondents reported no property crime with the remaining positives being quite spread with some extreme cases. Therefore, larger samples assist in estimating the coefficients of interest more precisely. To achieve the highest sample size possible, the sophisticated design of CJS is exploited, which apart from the core sample (representative sample of people between 10 and 65 years old), it also makes use of two 'boost' samples; the youth-boost sample (only young people between 10 and 25 years old) and the ethnic-boost sample (only non-white individuals between 10 and 65 years old). Each sample is accompanied by its (sampling) weighting variable. To re-establish representativeness, a weighting variable that combines the three separate weights is used.[11] A tabulation by sample type follows in Table 2.

- TABLE 2 ABOUT HERE -

The final data set used in the empirical investigation consists of 11,658 individuals, 5,604 males and 6,054 females, between 10 and 65 years old.[12] Notice that the sample size differs between the count and the binary form of the property crime variable. This is because some respondents 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 of the number of positive outcomes.

[11] 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 F of Hamlyn at al (2003).

[12] We need to note that by using individuals from 10 to 65 years old, we include students, retirees and people not in the labour force. Therefore, there is a departure from the economic model of crime. 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 household income and retirees receive a legal stream of pensions. In any case, the empirical analysis attempts to estimate the determinants of property crime for the whole population. This was inevitable, as limiting the sample only to individuals in the labour force reduces the sample size at a point where obtaining reliable results seemed impossible (at least for the models with under-reporting or misclassification).

Information about violent crime is also available and variables VIOLENT_B and VIOLENT_C are constructed (see Table 3).[13] Although this paper concentrates on property crime, results of violent crime are presented in subsection 6.2.3 to support the main findings.

- TABLE 3 ABOUT HERE -

The main explanatory variable of this study is the dummy IMMIGRANT which is 1 if someone is an immigrant and 0 if a native. This dummy is included in both reporting and crime equations. While it is common in empirical studies to define an immigrant as a person who is born outside the reference country, information about country of birth is not available in the questionnaire. 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 natives. Otherwise, they are classified as immigrants.[14] A limitation of this construction is that there can be some natives who left UK at some point, returning after a certain period of time. These people may have replied that they have lived in the UK less than their whole life, which makes them "immigrants", although they should be considered as natives, particularly if the period of staying outside the UK was very small. People who were born in the UK but lived most of their life in another country would exhibit very common characteristics with immigrants and therefore 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.[15] Although in the initial core sample only 729 immigrants appear, I have increased their number to around 2,000, by exploiting the youth-boost and most importantly the nonwhite-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.

Although the CJS provides a large set of respondents' characteristics, only controls for basic demographics, such as age, gender, region and ethnic background are considered in the empirical investigation. Of course, it would be interesting to explore the behaviour of the impact of immigration status on criminal behaviour or reporting behaviour once other controls such as income, education, working status, parental characteristics, perceived risks, drugs consumption etc, are included. However, because of mainly two reasons, it has been decided not to use this information in the empirical analysis. Firstly, 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

[13] Violent crime consists of: assaults with and without injuries, and commercial and personal robberies. Notice that both property and violent crime include robberies. However, also note that only 9 individuals reported a robbery.

[14] Respondents had to choose 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 but not the whole life; and 6) All of his/her life.

[15] 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.

2,500 individuals. Moreover, some other variables, such as risk factors, contain many missing cases which would reduce the sample size even more. 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.[16] Secondly, some of these variables such as education and perceived risks are to some extent endogenous in both crime and reporting equations, in the sense that they are correlated with unobserved respondents' characteristics that affect both criminal behaviour and reporting behaviour. Thus, conditioning on a set of endogenous covariates would generate biases, which would take very complicated forms since most of these covariates should appear in both equation. Moreover, since immigration status has a statistically significant effect on some of these variables, the immigration estimated coefficients will be biased in unknown directions. Instead, an "open" discussion will try to identify the factors that result in potential estimated crime differences between immigrants and natives. The remainder of this section provides a brief description of the covariates used in the empirical analysis (see, Table 3 for descriptives).

As mentioned in the previous paragraph, the following explanatory variables are used: AGE, is the age in years. MALE, takes value 1 if respondent is a male and 0 if female. Four dummies for the standard region of residence are used, which are, NORTH (1 if respondent lives in North, or Yorks & Humperside, or North West and 0 otherwise), MIDLANDS (1 if respondent lives in East Midlands, or West Midlands, or Wales and 0 otherwise), SOUTH (1 if respondent lives in East Anglia, or South East, or South West and 0 otherwise) and LONDON (1 if respondent lives in London and 0 otherwise) which will be the baseline group. Note that the grouping of regions into four groups was inevitable since estimation of the effects of all 9 dummies on both crime and reporting behaviour was impossible. Finally, five dummies to control for differences by ethnic groups are used, which are, WHITE, BLACK , ASIAN&OTHER and MIXED. Respondents that belong to ASIAN or OTHER ethnicity are grouped in one variable as 50% of people of OTHER ethnicity are Chinese. Note that, since the same individual is responsible for both actions of committing a crime and reporting a committed crime, logically both the crime and the reporting processes are functions of the same variables. Therefore, the variables described above will be included in both crime and reporting equations.

As described in Section 4.2, one strategy to identify the count data models is to find a variable that has no impact on the crime process, but a significant impact on the reporting process. The CJS provides some information which can be used to construct a variable that belongs to the reporting process only. Note that, as described in Section 4.1, the binary choice models are identified even without any exclusion restriction. However, this extra information will be also used in binary choice models both because we would like the analysis to be consistent across all models, but also because exclusion restrictions can facilitate the estimation procedure. However, as will be made clear in subsection 7.4 the exclusion restriction does not drive the results of the binary models. In the next two paragraphs two available options are described.

[16] Actually, to achieve convergence we had to impose several exclusion restrictions from both processes (crime and reporting), which leads to serious model misspecification, since some of the excluded variables actually belong to both processes.

Firstly, respondents were asked whether they replied to the questions related to crime truthfully. Thus, the dummy variable TRUTHFUL is generated which takes value 1 if respondents said that they were truthful and 0 otherwise. This variable is used only in the reporting process, as whether or not someone truthfully reported his/her actual criminal activity at the time of the survey cannot have affected criminal activity prior to the survey. If any empirical relationship exists, this would be because "truthfulness" is correlated with unobserved characteristics which affect criminal behaviour, or because there is reverse causality of committed crimes on "truthfulness".[17] However, at the same time, it is not appropriate to assume that "truthfulness" actually affects the reporting behaviour, unless the reported "truthfulness" coincides with the actual behavioural characteristic of how truthful someone is. However, what we assume here is that being "truthful" while answering questions related to crime is a feature that "shapes" some behavioural attributes, which in turn affect reporting behaviour. In any way, when TRUTHFUL is included in both processes, the results show that it actually has a significant impact on the reporting process but no effect on the crime process.

Secondly, in 32% of the interviews (3,768 observations) there was someone else present during the interview, mostly in the cases of young individuals. 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 behaviour. Therefore, the dummy OTHER_PRESENT is constructed that takes value 1 if someone else was present and 0 otherwise. However, the results show, that this variable has no effect on the reporting process, making identification of the parameters of the count data models very difficult. This can be attributed to the fact that not only were crime questions self-completed in a computer, but also because it was stressed by the interviewers that nobody should disturb the interviewee during the self-completion part.[18] Therefore, the main results are obtained exploiting the variable TRUTHFUL only. However, as we will see in subsection 7.4, the use of OTHER_PRESENT provides some interesting insights.

6 Main Results

The main Probit results, using PROPERTY_B as dependent variable, are presented in Table 4, in 4 specifications. Specification (1) presents the profit estimates without controlling for misreporting ($\tilde{\boldsymbol{\beta}}$), while specifications (2), (3) and (4) present the MisProbit results, where $\hat{\boldsymbol{\beta}}$ are the estimated effects of the regressors on the probability to commit a property crime and $\hat{\boldsymbol{\gamma}}$ are the estimated effects on under-reporting. We also include the square of variable AGE to capture a potential

[17] 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 likely 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.

[18] Note that we also have information on whether "someone looked at the screen during the self completion part" which is used to construct the dummy variable LOOKED_SCREEN. However, the results show that this dummy has no effect either using it as an interaction term OTHER_PRESENT×LOOKED_SCREEN or using it alone without controlling for the cases where someone was present but did not look at the screen.

quadratic relationship between age and criminal behaviour, or age and reporting behaviour.

- Table 4 About Here -

We begin with the results of models that do not control for under-reporting. Table 3 shows that without including any control variables, being an immigrant decreases the probability of reporting a crime by 1.8 percentage points, an effect that is significant at 5% significance level. This difference is quite large in magnitude if we consider that only 5.5% of respondents reported a property crime. The results of specification (1) in Table 4, however, show that although the sign remains negative, the effect of immigration status becomes insignificant once we control for basic demographics. Furthermore, if we control for ethnicity the estimated effect gets even closer to zero. This is because immigrants are more nonwhite relative to natives but white individuals report more crime than nonwhite ones.[19]

Of course, the results discussed above are not very informative, since the estimated differences may reflect either a true difference in criminal activity, or merely differences in reporting behaviour. For this reason, we proceed with the MisProbit results which are presented in specifications (2), (3) and (4) of Table 4. Before discussing the main findings, however, there are two important issues we need to stress. First, notice that while under-reporting is allowed to depend on covariates, over-reporting is assumed to be constant.[20] Second, an exclusion restriction on the crime process is used as discussed in the previous section. Here, we assume that variable Truthful affects the reporting process but not the crime process.

At these stage, it is important to mention some main features of the findings of this model. First of all, these models predict that the average probability of misclassification of one as zero, calculated as $\hat{\alpha}_0 = \sum_{i=1}^{n} \Phi(z_i'\hat{\gamma})/n$ is around 0.70. This seems fairly large, as it indicates that around 70% of the population commits crimes but reports none of them. However, as emphasised in Section 4, 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). 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. The estimated value of this probability $\hat{\alpha}_1 = 0.012$ is also expected, as we would not expect people to report crimes they have not committed. Note, that both $\hat{\alpha}_0$ and $\hat{\alpha}_1$ are significant at 1% level of significance. Moreover, the average probability of committing a property crime, calculated as $\widehat{\Pr(y_i^* = 1)} = \sum_{i=1}^{n} \Phi(\mathbf{x}_i'\hat{\boldsymbol{\beta}})/n$, is around 0.29, which is much higher than the predicted average probability of the simple Probit model, calculated to be 0.064. However, notice that if we accept the interpretation of misclassification

[19] Note also that the Age and Age^2 do not provide a good fit. To obtain the pattern of the impact of age on the observed property crime we need to include up to the fourth power of age. Following this, the effect of immigration dummy becomes slightly more significant but it remains insignificant. Here, we present the model that uses Age^2 only in order to be in line with the specifications we use in the models that control for misreporting.

[20] 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 regressors. Nevertheless, the estimation analysis showed that identifying all coefficients of a 'fully' specified model seemed impossible.

of 1 as 0 as zero-inflation, this is actually the predicted probability of committing a crime only for those that participate in the binary choice model. Finally, as expected, the standard errors of the MisProbit are fairly larger than the simple Probit model. Nevertheless, provided that the data are truly generated by the MisProbit model, these standard errors provide estimates of the true uncertainty, as the standard errors of the Probit model are actually overestimated.

Concerning the main objective of this research work, the results show that after controlling for potential differences in the reporting behaviour between immigrants and natives, the effect of immigration status on properly crime is still negative and larger in magnitude than the Probit model. In specification (2) it is actually statistically significant at 10% level. However, after controlling for ethnicity, although still negative and fairly large, it becomes insignificant. The partial effect of being an immigrant, calculated for a "representative" individual, who is a 25 year-old male that lives in London, tells us that being an 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.[21] The reason why the estimate on immigration status becomes larger in magnitude may be attributed to the negative effect that being an immigrant has on the reporting process, even though insignificant, indicating that native-born individuals in fact under-report by more than immigrants. However, since the coefficients of the reporting process can also take a zero-inflation interpretation, the negative coefficient might also mean that a smaller proportion of immigrants belong to the group of genuine non-criminals.

It would be also interesting to briefly discuss the effects of the other explanatory variables. To begin with, it is noteworthy that being a white individual increases the probability of committing a crime, an effect which is significant at 10% level. This difference is driven by the negative effects of BLACK and ASIAN&OTHERS variables. Regarding gender, as expected, being a male increases the probability of committing a property crime, an effect that is significant at 1% significance level. Interestingly, the effect of being male on the reporting process is negative, indicating that females are more reluctant to report their criminal activities truthfully, perhaps because of "embarrassment" effects. However, this may also indicate that females are more likely to belong to the genuine non criminal group of people, which is also reasonable. AGE seems to have a quadratic significant effect on both crime and under-reporting, where both crime and under-porting decreases with age in a decreasing rate. A quick calculation shows that the probabilities to commit a crime and to under-report a committed crime are minimised at about 42 and 33 years of age, respectively. Note that experimenting with different powers of age does not provide a better fit. The results of the regional dummies suggest that people who do not live in London are more likely to commit a property crime. Nevertheless, only people living in North England seem to commit significantly more crime (but just in 10% level). Finally, the effect of being TRUTHFUL is negative and strongly significant. This indicates that either, respondents who said that they answered all crime questions truthfully are honest people who, regardless of their

[21] The predicted probabilities to commit a crime for the "representative" individual are 4.27% for an immigrant, and 9.87% for a native (2.3 times higher for a native). This corresponds to a partial effect of 5.6 percentage points. After controlling for ethnicity the above figures become $\approx$ 6%, 10% and 4 percentage points respectively.

observed characteristics, never commit and never report crimes, or that "truthful" respondents in fact under-report by more than "non-truthful" respondents.

7 Robustness Checks

7.1 Count Data Models

In this section we examine whether the results of count data regression models are in line with the main findings of the binary models. Of course, these two are not directly comparable, since the count data models model the conditional mean of crime events, by using the extra information of the number of property crimes, whereas MisProbit models the conditional probability of committing a crime. Therefore, even if both models are correctly specified, it is always possible that a few differences exist between binary and count data models. Nevertheless, similar estimates, mostly with respect to the reporting process, would strengthen the reliability of the results of the main analysis. Specifications (1), (2) and (3) of Table 5 present the results of the NB2, NB2-Logit and ZI-NB2-Logit, respectively.

- TABLE 5 ABOUT HERE -

Note that, in line with the binary choice models, the TRUTHFUL dummy is used in the reporting process only. However, the exclusion restriction is much more crucial here, as otherwise NB2-Logit and ZI-NB2-Logit are not identified. However, although the two models are globally identified since TRUTHFUL is a "strong" exclusion restriction, we must still be very cautious since more than one maxima exist.[22] Although our investigations showed that several maxima exist, the global maxima are the ones presented in Table 5.[23] As explained before, these models do not control for over-reporting, differing in this aspect from binary choice models. However, we find that MisProbit gives very similar results even when the probability of over-reporting is assumed to be zero.[24] 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 model only provide information about total under-reporting, regardless of how many crimes someone has committed. The count data models on the other hand, provide estimates for the probability of any given committed crime to be reported.[25]

Regarding the estimates of the NB2-Logit model, first of all, the large estimated value of α must be noticed, which is statistical significant at any significance level.[26] Therefore, there is

[22]Some tips to find the best maximum are described in Papadopoulos and Santos Silva (2008).

[23]The estimation analysis showed two local maxima very close to the ones presented in Table 5 exist with log likelihood values of -2,315.80 and -2,261.28 for the NB2-Logit and ZI-NB2-Logit respectively. These maxima correspond to estimates that are very different from the ones that correspond to the global maxima.

[24]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.

[25]Although this probability is allowed to differ across individuals, it is assumed to be constant for all committed crimes by individual i regardless of the number of committed crimes he/she has committed.

[26]Note, that the NB2-Logit reduces to the Poisson-Logit if $\alpha = 0$ (see, Wnkelmann and Zimmermann, 1993). Note also that the log likelihood value of the corresponding global Poisson-Logit maximum is -9,132.31.

evidence that the data are over-dispersed even after conditioning on the set of regressors. As far as the reporting process is concerned, this model predicts that the average conditional probability of reporting a committed crime, calculated as $\hat{p} = \sum_{i=1}^{n} \Lambda(\mathbf{z}_i'\hat{\boldsymbol{\eta}})/n$, is 43%. We can also see that, as expected, most of the coefficients of the Logit part have opposite signs to the ones of the under-reporting part of the MisProbit model.

Concerning the crime process, we can first notice that the signs of estimates $\hat{\boldsymbol{\delta}}$ are in line with the MisProbit's estimates $\hat{\boldsymbol{\beta}}$. Moreover, even after controlling for under-reporting, the effect of being an immigrant on the expected number of crimes is negative and even larger in magnitude than the conventional NB2 model. This is because being an immigrant increases the probability of reporting a committed crime and therefore decreases the conditional expected number of crimes by more than the conventional NB2 model. However, in accordance with the binary models, the coefficient of immigration dummy is insignificant in both processes.[27] Finally, the partial effect of our "representative" individual says that being an immigrant decreases the expected number of crimes by around 0.19, which difference was 0.11 for the simple NB2 model.[28]

The estimates of the ZI-NB2-Logit model are presented in specification (3). Since this model allows for a probability of zero-inflation, the Logit process measures the probability of reporting a committed crime once we partial out people who always under-report with zero, or people that never commit and consequently never report crimes. According to this model, once we control for zero-inflation and under-reporting, the estimated effect of immigration status becomes even larger in magnitude, but still statistically insignificant as the precision of the estimates decreases. As the "zero-inflation" process of the ZI-NB2-Logit has the same interpretation as the "misclassification of one as zero" process of the MisProbit, it would be interesting to compare the corresponding coefficients of the two models. Firstly, we notice that apart from the coefficients on Age variables, all other coefficients follow the same direction, although there are differences in the statistical significance of some coefficients. Moreover, we interestingly find that the predicted average conditional probability of zero-inflation, calculated as $\hat{\xi} = \sum_{i=1}^{n} \Lambda(\mathbf{q}_i'\hat{\boldsymbol{\vartheta}})/n$ is around 62%, a figure that is close to what the MisProbit model predicts. Finally, the coefficients of the "reporting a committed crime" process are relatively similar across the two NB2 models.

7.2 Violent Crime

In this subsection we briefly investigate the link between immigration status and violent crime. The violent crime estimates, using the specification (3) of Table 4, are presented in specification (1) of Table 6. Impressively, the estimates of property and violent crime are remarkably close, apart from the coefficients of the regional dummies.[29] We can see that the same basic demo-

[27]Note that in NB2-Logit, contrary to MisProbit, the effect of immigration dummy is insignificant even when we do not control for ethnicity.

[28]The expected number of committed crimes by the "representative" individual are, 0.1771, and 0.3628, for an immigrant and a native, respectively, 2 times larger for a native.

[29]Note that 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

graphical characteristics are good predictors of both violent and property crime. In addition, we can notice that the predicted probability of committing a violent crime but not reporting it is lower than for property crime. Concerning the effect of the immigration dummy, it is again negative but slightly less significant.

- TABLE 6 ABOUT HERE -

7.3 Weighted *versus* Unweighted Regressions of Property Crime

The presented estimates so far are obtained utilising regression models that make use of the appropriate weights to restore representativeness of the sample. However, if the conditional expectation is correctly specified, both weighted and unweighted estimators are consistent, but the unweighted one is also more efficient (see, Wooldridge, 2010). 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. The estimates are presented in specification (1) of Table 7 and specification (2) of Table 6, for property and violent crime respectively.

- TABLE 7 ABOUT HERE -

It is noteworthy, that the estimated coefficients of the weighted estimates of the MisProbit are very close to the unweighted ones for both property and violent crime. Moreover, it is evident that the coefficients of the unweighted regression are more precisely estimated. A notable difference however is that in the unweighted estimation the effect of immigration is higher, in terms of magnitude, and statistically significant at 5% for property crime and 10% for violent crime. 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. Note also that as we have included an ethnic-boost data set, immigrant population is over-represented in the sample. Thus, using weights has as a result to down-weigh the immigration sample, which results in less precise estimates. Finally, notice that the biggest differences are observed in the estimated coefficients of the variables that are insignificant when we use weights, such as the regional dummies' ones.

7.4 Are the Results Driven by the Exclusion Restriction?

In this subsection I briefly intend to show that the main results are not driven by the exclusion of variable TRUTHFUL from the crime process. Indeed, the estimates of the MisProbit without any exclusion restrictions, are quite similar to those which use variable TRUTHFUL both for property and violent crimes (see, specification (2) of Table 7 and specification (3) for property and violent crime respectively). 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.

crime and 1.1% for violence.

In specification (3) of Table 7, we look at the consequences of including OTHER_PRESENT in the reporting process only, where we notice that this dummy has no effect on the reporting process of property crime. Notice also that the inclusion of OTHER_PRESENT actually results in much less precise estimates for most of the parameters. 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 behaviour of the model. Consequently, the effect of being an immigrant becomes more insignificant. On the other hand, it occurs that OTHER_PRESENT has a significant positive effect (at 1% significance level) on the the reporting process of violent crime (see, specification (4), Table 6). Contrary to property crime, it is noteworthy that the estimates of (4) are very close to the estimates when no exclusion, or the TRUTHFUL exclusion are used. Again, the magnitude of immigrants' coefficient slightly decreases but so does the standard error, leaving significance almost unaffected. Thus, overall, there is some evidence that the MisProbit results are robust in relation to the exclusion restriction, as long as the exclusion restriction is "strong".

Regarding count data models, specifications (4) and (5) of Table 5 present results of including the dummy OTHER_PRESENT in the reporting process of NB2-Logit to test whether the results of count data models are also robust in relation to the exclusion restriction.[30] As can be seen from (4), OTHER_PRESENT has again no effect on the probability of reporting a committed property crime. However, in this case another maximum very close (in terms of the log likelihood value) to the global maximum exists that corresponds to very different parameter estimates. The estimates of the second maximum are presented in specification (5). As Papadopoulos and Santos Silva (2008) show, which is also very clear in this application, it appears that there is a close relationship between the parameters of the two maxima. Given that $\boldsymbol{\theta} = (\boldsymbol{\delta}, \boldsymbol{\eta})$ is the set of true parameters of the model, if the exclusion restriction is "weak", another maximum very close to the true one exists with parameter values $\tilde{\boldsymbol{\theta}} \simeq (\boldsymbol{\delta} + \boldsymbol{\eta}, -\boldsymbol{\eta})$. 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 $\tilde{\boldsymbol{\theta}}$ from $(\boldsymbol{\delta} + \boldsymbol{\eta}, -\boldsymbol{\eta})$. Despite the identification problem of this model, assuming that the correct maximum is the one in (4), it is clear that the estimated parameters are very similar whether we use TRUTHFUL or OTHER_PRESENT as an exclusion restriction.

7.5 More Robustness Checks

Further robustness checks were conducted, whose results are not presented here but are available on request. First, we checked the consequences of dropping from the sample all 117 immigrants who reported that have been in the country for less than 12 months. This is a potentially interesting exercise because of two reasons: firstly, the CJS does not record crimes that happened outside the UK. Since the crime questions concern individuals' criminal behaviour during the 12 months prior to the day of the interview, there might be some cases of very recent immigrants who committed crimes outside the UK which are not recorded. However, at the same time,

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

some of the most recent immigrants may have committed crimes in their source countries and mistakenly recorded them as if they happened in the UK. Nevertheless, we would not like to include these reported crimes in our sample either, since their countries of origin may exhibit very different characteristics associated with property crime, such as different economic opportunities and deterrent factors, resulting in a different criminal behaviour. Therefore, by dropping these 117 cases we avoid these two ambiguous scenarios. The results show that by doing this, the effect of immigration status slightly increases in magnitude but it is still statistically insignificant.

In addition we looked at the consequences of dropping very young individuals (10 to 13 year olds), as responses of children might be less reliable. Notice, however, that because we use the youth-boost sample, dropping very young individuals results in losing many observations decreasing precision of our estimates. The estimates indicated that excluding young respondents, the effect of being an immigrant, although less precise, becomes larger in magnitude and actually turns significant at 10% if we exclude respondents up to 13 years old. Finally, note that if we drop 14 year-old individuals, which reduces the sample size by 1,785 observations, results in no convergence of the estimation procedure.

Finally, note that even though criminal damage is also a crime against the property, it entails only psychological gains to the offenders and therefore, it is not very clear whether it is proper to include it in the property crime variable. This is because, as it is the case for violent crime, criminal damage cannot be well explained by the economic model of crime. The results suggest that, excluding criminal damage which increases the proportion of zeroes from 0.945 to 0.951, the effect of immigration status on property crime slightly reduces in magnitude but it still retains its sign.

8 Interaction Terms between Immigration Status and Region, or, Ethnicity

In this subsection, we investigate whether the effect of immigration status on property crime depends on the region of residence or on ethnicity using interaction terms. The results are presented in Table 8 and are briefly discussed.

- Table 8 About Here -

To begin with, since location of immigrants is not randomly assigned, different locations may attract different criminal-types of immigrants, or, immigrants located in different places may face different conditions, which in turn may affect their criminal behaviour. From specification (1) of Table 8 and using appropriate Wald tests, there are two things that merit some discussion. Firstly, although as a whole immigrants are not significantly less involved in criminal activities, it is interesting that immigrants located in London are significantly (at 1 % level) less likely to commit property crimes than Natives in London. Actually, immigrant population in London is the least crime-prone group, since they exhibit a significantly (at 1 % level) lower involvement in criminal activities, not only compared to all groups of natives but also compared to immigrants that are located in South. On the other hand, it is also interesting that immigrant population located in South is the most crime-prone category. However, their involvement in crime is not statistically different from the involvement of natives in South. Finally, note that natives exhibit similar criminal behaviour across all regions and that immigrants in Midlands and North are not significantly less likely to commit property crimes than their native counterparts.

But which possible channels could explain the results above? It might be, for instance, that immigrants integrate in London more easily than in other locations, because of better labour opportunities and a larger concentration of immigrants. At the same time, this high concentration of immigrants in specific areas of London might generate strong social controls that discourage criminal activities. In addition, if immigrants are more responsive to deterrent factors (see, for example, Bucher and Piehl, 2007), strict policing in London would discourage criminal activities of immigrants by more than natives. 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 and Fasani, 2010). On the other hand, immigrants located in South may encounter problems of adaptation in the English society, or the socio-economic conditions they face may be less favourable than those of other regions. Finally, perhaps South pulls the most crime-prone groups of immigrants, simply because the risk of apprehension may be lower in South than other regions, as escape to other countries in continental Europe in a case of a legal issue seems easier.

As a second exercise we examine whether the immigration-crime relationship differs among different ethnic groups. This might be the case since immigrants of different ethnic status may have grown up in environments with quite different principles and values, or, in different socio-economic conditions. In addition, we might also expect that immigrants of one ethnic group may exhibit different criminal behaviour from natives of the same ethnic group, as the latter is better adapted in the British lifestyle.

First of all, comparing each ethnic group of immigrants with the native population as a whole (that is, regardless of natives' ethnicity), we find that being a Black immigrant significantly (at 1 % level) decreases the probability to commit a property crime. Moreover, Asian&Other immigrants are also less crime-prone than natives, but the difference is significant only at 20% level.[31] Moreover, using the results in specification (2) of Table 8 and appropriate Wald tests, the most impressive result we find is that Black immigrants are also less likely to commit a property crime than Black natives (a difference that is actually significant at 5% level). Thus, Black immigrants is in fact the least crime-prone group, which is very interesting if we consider that this also the group, particularly those coming from Africa, that faces the most unfavourable socio-economic conditions (see, for example, Algan at al, 2010). Note also that the involvement of Black natives in criminal activities is not different than the involvement of all other groups. Thus,

[31]Note that these results are note presented in Table 8, but are available on request.

Black immigrants exhibit unobserved cultural characteristics associated with lower involvement in criminal activity than the other groups. Finally, note that there is no difference in crime between the other three immigrant groups and their native counterparts.

9 Concluding Remarks

This study investigated the individual relationship between immigration and property crime in England and Wales using crime self-reports from the CJS. Since under-reporting is a major concern in crime self-reports, regression models for binary and count data that control for this problem were used. Given that the assumptions of these models hold, we are able to consistently estimate the determinants of true criminal activity and the determinant of misreporting using only data of reported crime.

Although there is a public sentiment that immigrants are more involved in criminal activities, the results of these models lead to different conclusions. Firstly, although there is evidence of significant under-reporting of criminal activity, if anything, immigrants tend to under-report by less than natives. Nevertheless, it was stressed that the coefficients of the reporting process of the MisProbit model must be treated with caution, since the reporting process can be also interpreted in a zero-inflation framework. That is, the model cannot distinguish between the probability of reporting zero crimes because of under-reporting and reporting zero crimes because of never committing and consequently never reporting any crimes. The estimates indicated that reporting zero crimes because of total under-reporting or because of zero-inflation, conditional on the set of covariates, is around 70%.

The estimates of the crime process suggested that, controlling for differences in reporting behaviour, if immigrants were similar to natives in terms of basic demographic characteristics, there would be a negative association between actual criminal behaviour and immigration status. Even though the estimated difference is statistically insignificant in most specifications, all the results in the sensitivity analysis section implied that it is actually quite robust. For example, the results of the unweighted MisProbit signified that if we were able to obtain a larger sample, the estimated negative association would be much more precise. Therefore, altogether, the robustness of the association might suggest that this relationship actually exists in the population, but the nature of the regression models in combination with the data in hand do not allow estimating the relationship more precisely.

The economic model of crime presented in Section 2, showed that even though there are several channels through which immigration can be associated with crime, the sign of this association is not clear. How can the immigration-crime estimates 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, because employers tend to prefer natives, etc (see, Algan at al, 2010). There are also, to some extent, cultural conflicts, and difficulties of adjustment. However, at the same time, immigrants may be more risk averse and discount future less heavily. As a result, they might be more responsive to potential punishment and other deterrent factors (Bucher and Piehl, 2007). In addition, not only do immigrants face a higher probability of apprehension, but they are also confronted with the threat of deportation. 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 compared to natives.

Finally we showed, using interaction terms, that the effect of immigration status on property crime actually depends on the region of residence and ethnicity. 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. Thus, it might be that either, different socio-economic conditions that immigrants encounter in different locations and their interactions with the native population may affect their criminal behaviour, or that different areas attract different types of immigrants. Finally, we interestingly found that, because unobserved cultural factors, Black immigrants are more crime-averse than black natives and white natives, despite the fact that they are the least favoured group with regard to their socio-economic characteristics. However, we need to stress that the analysis of interaction terms is limited by the small sample size and the very small number of positive outcomes associated with each separate group. Further investigation is required to establish whether the effect of being an immigrant on criminal behaviour differs with respect to immigrants' demographic characteristics.

References

Albrecht, H. (1997), "Ethnic Minorities, Crime, and Criminal Justice in Germany", in M. 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.

Alesina, A., and E. La Ferrara (2005), "Ethnic Diversity and Economic Performance", *Journal of Economic Literature*, 43, 762-800.

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

Beine, M., F. Docquier, and H. Rapoport (2008), "Brain Drain and Human Capital Formation in Developing Countries", *Economic Journal*, 118, 631-652.

Bell, B. and S. Machin (2011), "Immigrant Enclaves and Crime", IZA Discussion Paper No 6205.

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.

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

Borjas, G. J. (1999), "Immigration and Welfare Magnets", *Journal of Labor Economics*, 17, 607-637.

Borjas, G. J. (2003), "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market", *Quarterly Journal of Economics*, 118, 1335-1374.

Bucher, F. K., and A. M. Piehl (1998), "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 (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 Research Study 275.

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

Card, D (2009), "Immigration and Inequality", *American Economic Review*, 99, 1-21.

Chaplin R., J. Flatley, and K. Smith (2011), "Crime in England and Wales 2010/11", Home Office Statistical Bulletin, 10/11.

Dustman, C., F. Fabbri, and I. Preston (2005), "The Impact of Immigration on the British Labour Market", *Economic Journal*, 115, F324-F341.

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.

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

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

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

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

Hart, H. H. (1896), "Immigration and Crime", *American Journal of Sociology*, 2, 369-377.

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

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

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

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

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. Marshall (1999), "The Self-Report Methodology in Crime Research", *Crime & Justice*, 291-367.

Killias, M. (1997), "Immigrants, Crime, and Criminal Justice in Switzerland", in M. 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.

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

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.

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

Martens, P. L. (1997), "Immigrants, Crime, and Criminal Justice in Sweden, in M. 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.

Manacorda, M., A. Manning and J. Wadsworth, (2012), "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain", *Journal of the European Economic Association*, 10(, 120-151.

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. (2011), "The Relationship between Immigration Status and Criminal Behaviour: Evidence from the Offending, Crime, and Justice Survey", Unpublished PhD Manuscript, Department of Economics, University of Essex.

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.

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.

Smith, D. J. (1997), "Ethnic Origins, Crime, and Criminal Justice in England and Wales", in M. 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.

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

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.

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

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.

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

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

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

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

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

Table 1: Distribution of Property Crime

Num of Property Crimes	Count	Percent
0	10,927	94.17
1	251	2.16
2	123	1.06
3	67	0.58
4	40	0.34
5	48	0.41
6-10	63	0.54
11-15	31	0.27
16-25	25	0.22
26-49	16	0.14
50-99	7	0.06
100-225	6	0.05
Total	11,604	100.00

Table 2: Tabulation of CJS Respondents by Sample Type

Sample Type	Total		Immigrants		Natives	
	N	%	N	%	N	%
Core 10-65	6,771	58.08	729	36.25	6,042	62.63
Youth Boost 10-25	3,098	26.57	186	9.25	2,912	30.19
Ethnic Boost 10-65	1,789	15.35	1,096	54.50	693	7.18
Total	11,658	100	2,011	100	9,647	100

Table 3: Descriptive Statistics (Weighted Means by Immigrations Status)

Variables	Weighted Mean			Min	Max
	All	Natives	Immigrants		
Crime Variables					
PROPERTY_B	0.055	0.057	0.039	0	1
VIOLENT_B	0.054	0.056	0.036	0	1
PROPERTY_C	0.342	0.366	0.160	0	225
	(5.858)	(6.011)	(2.218)		
VIOLENT_C	0.286	0.310	0.107	0	512
	(4.132)	(4.252)	(0.918)		
Explanatory Variables					
IMMIGRANT	0.119			0	1
AGE	36.738	36.554	38.098	10	66
MALE	0.497	0.496	0.505	0	1
WHITE	0.909	0.956	0.553	0	1
BLACK	0.023	0.010	0.120	0	1
ASIAN&OTHER	0.056	0.022	0.214	0	1
MIXED	0.012	0.008	0.039	0	1
NORTH	0.274	0.288	0.175	0	1
MIDLANDS	0.235	0.246	0.150	0	1
SOUTH	0.351	0.358	0.298	0	1
LONDON	0.139	0.107	0.376	0	1
TRUTHFUL	0.942	0.946	0.915	0	1
OTHER_PRESENT	0.285	0.288	0.263	0	1

Weighted Standard Deviations in Parentheses

Table 4: Probit Results

PROPERTY_B	Probit (1) $\tilde{\boldsymbol{\beta}}$	MisProbit (2) $\hat{\boldsymbol{\beta}}$	(2) $\hat{\gamma}$	MisProbit (3) $\hat{\boldsymbol{\beta}}$	(3) $\hat{\gamma}$	MisProbit (4) $\hat{\boldsymbol{\beta}}$	(4) $\hat{\gamma}$
IMMIGRANT	-0.127	-0.431*	-0.391	-0.254	-0.410	-0.277	-0.437
	(0.102)	(0.251)	(0.371)	(0.273)	(0.346)	(0.275)	(0.359)
AGE	-.0178**	-0.248***	-0.205**	-0.259***	-0.195***	-0.255***	-0.193***
	(0.007)	(0.045)	(0.081)	(0.046)	(0.074)	(0.047)	(0.074)
AGE^2	0.0001	0.003***	0.003**	0.003***	0.003***	0.003***	0.003***
	(0.0001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
MALE	0.362***	0.439***	-0.260*	0.440***	-0.277*	0.429***	-0.285**
	(0.049)	(0.156)	(0.144)	(0.156)	(0.142)	(0.156)	(0.142)
WHITE				0.542*	-0.198		
				(0.293)	(0.224)		
BLACK						-0.559*	-0.264
						(0.323)	(0.447)
ASIAN&OTHER						-0.638	0.467
						(0.430)	(0.327)
MIXED						0.004	0.104
						(0.391)	(0.281)
SOUTH	0.089	0.300	0.201	0.273	0.262	0.272	0.252
	(0.082)	(0.219)	(0.228)	(0.231)	(0.224)	(0.227)	(0.221)
MIDLANDS	0.061	0.151	0.151	0.124	0.224	0.124	0.211
	(0.084)	(0.209)	(0.215)	(0.228)	(0.215)	(0.223)	(0.212)
NORTH	0.066	0.501*	0.432**	0.496	0.490**	0.498	0.480**
	(0.086)	(0.282)	(0.211)	(0.318)	(0.212)	(0.317)	(0.207)
TRUTHFUL			0.876***		0.868***		0.854***
			(0.319)		(0.256)		(0.255)
CONSTANT	-1.204***	2.597***	2.196***	2.335***	2.267***	2.833***	2.072***
	(0.127)	(0.895)	(0.522)	(0.896)	(0.507)	(0.926)	(0.516)
Sample Size	11,658	11,658		11,658		11,658	
Log Likelihood	-1452.88	-1427.97		-1422.05		-1419.99	
$\widehat{\Pr(y^* = 1)}$	0.064	0.285		0.292		0.295	
$\hat{a}_0$		0.681		0.709		0.709	
$\hat{a}_1$		0.012		0.013		0.012	

Quasi Maximum Likelihood Standard Errors in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively.

Table 5: Negative Binomial Results

PROPERTY_C	NB2	NB2-Logit		ZI-NB2-Logit			NB2-Logit (Other Present)			
	(1)	(2)			(3)		(4)		(5)	
	$\tilde{\delta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\hat{\delta}}$	$\hat{\hat{\eta}}$	$\hat{\hat{\vartheta}}$	$\hat{\delta}$	$\hat{\eta}$	$\check{\delta}$	$\breve{\eta}$
IMMIGRANT	-0.309	-0.617	0.451	-0.757	0.130	-0.533	-0.475	0.189	-0.306	-0.094
	(0.352)	(0.636)	(1.043)	(0.721)	(0.869)	(0.549)	(0.595)	(0.972)	(0.640)	(0.931)
AGE	0.002	-0.677***	1.008***	- 0.711**	1.058***	0.138**	-0.619***	0.910***	0.422***	-0.939***
	(0.030)	(0.224)	(0.206)	(0.308)	(0.230)	(0.060)	(0.180)	(0.208)	(0.145)	(0.182)
AGE^2	-0.001*	0.009***	-0.015***	0.010**	-0.015***	-0.001	0.009***	-0.014***	-0.007***	0.014***
	(0.000)	(0.003)	(0.003)	(0.005)	(0.004)	(0.001)	(0.003)	(0.003)	(0.002)	(0.003)
MALE	0.672**	1.457***	-1.333*	0.657	-0.406	-0.960***	1.462***	-1.325**	0.125	1.139
	(0.275)	(0.443)	(0.687)	(0.498)	(0.625)	(0.259)	(0.394)	(0.659)	(0.477)	(0.708)
WHITE	0.788***	0.196	0.638	-0.939	1.053	-1.634***	0.175	0.734	0.869	-0.525
	(0.268)	(0.611)	(1.062)	(0.706)	(0.994)	(0.387)	(0.662)	(1.191)	(0.611)	(0.974)
SOUTH	0.305	1.024	-1.508	1.504**	-1.890**	0.49	0.860	-1.024	-0.270**	0.992
	(0.279)	(0.696)	(0.937)	(0.694)	(0.771)	(0.480)	(0.601)	(0.902)	(0.507)	(0.771)
MIDLANDS	0.441	0.247	0.255	0.501	-0.48	0.136	0.161	0.519	0.694	-0.416
	(0.269)	(0.561)	(0.954)	(0.682)	(0.878)	(0.561)	(0.558)	(0.995)	(0.627)	(0.878)
NORTH	1.065 **	1.676**	-1.628	2.733***	-2.800***	0.820*	1.588**	-1.373	-0.201	1.903
	(0.507)	(0.662)	(1.027)	(0.949)	(1.010)	(0.479)	(0.711)	(1.164)	(0.575)	(0.845)
TRUTHFUL			-1.237***		-1.963***					
			(0.475)		(0.468)					
OTHER_PRESENT								-0.677		-0.600*
								(0.533)		(0.349)
CONSTANT	-1.897***	8.633**	-12.253***	10.950**	-13.285***	-1.212	7.572**	-11.470***	-5.849***	11.870***
	(0.529)	(3.671)	(3.510)	(5.288)	(4.520)	(1.133)	(3.009)	(3.664)	(1.538)	(2.799)
Sample Size	11,604	(11,604)			11,604		11,604		11,604	
Log Likelihood	-2,344.45	-2,313.99			-2,258.49		-2,314.29		-2,314.61	
$\hat{\alpha}$	46.66***	41.64***			15.474***		41.94***		41.99***	
$\hat{p}$		0.430			0.376		0.448		0.464	
$\hat{\xi}$					0.617					

Quasi Maximum Likelihood Standard Errors in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively.

Table 6: Robustness Checks - Violent Crime

VIOLENT_B	Truthful (1)		No Weights (2)		No Exclusion (3)		Other Present (4)	
	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$
IMMIGRANT	-0.241	-0.553	-0.378*	-0.258	-0.177	-0.502	-0.169	-0.503
	(0.341)	(0.416)	(0.203)	(0.268)	(0.271)	(0.423)	(0.259)	(0.440)
AGE	-0.315***	-0.646***	-0.347***	-0.562***	-0.296***	-0.723***	-0.291***	-0.742***
	(0.056)	(0.154)	(0.060)	(0.148)	(0.044)	(0.110)	(0.040)	(0.135)
AGE^2	0.004***	0.012***	0.005***	0.011***	0.004***	0.014***	0.004***	0.014***
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.004)
MALE	0.428***	-0.229	0.295***	-0.302**	0.410***	-0.283*	0.419***	-0.276*
	(0.102)	(0.149)	(0.094)	(0.117)	(0.093)	(0.150)	(0.092)	(0.157)
WHITE	0.447***	-0.057	0.118	-0.316	0.429***	-0.060	0.447***	-0.071
	(0.172)	(0.264)	(0.147)	(0.183)	(0.158)	(0.266)	(0.172)	(0.269)
SOUTH	-0.063	-0.237	-0.040	-0.109	-0.050	-0.186	-0.042	-0.165
	(0.241)	(0.310)	(0.164)	(0.205)	(0.213)	(0.319)	(0.212)	(0.337)
MIDLANDS	-0.085	-0.063	-0.044	0.038	-0.085	-0.050	-0.074	-0.020
	(0.242)	(0.309)	(0.169)	(0.215)	(0.216)	(0.321)	(0.210)	(0.343)
NORTH	-0.140	-0.256	-0.120	-0.134	-0.128	-0.259	-0.117	-0.252
	(0.246)	(0.316)	(0.173)	(0.212)	(0.158)	(0.327)	(0.211)	(0.330)
TRUTHFUL		0.876***		0.777***				
		(0.253)		(0.199)				
OTHER_PRESENT								0.350***
								(0.160)
CONSTANT	2.778**	6.174***	3.667***	5.720***	2.470***	7.666***	2.354***	7.509***
	(0.925)	(1.250)	(1.048)	(1.184)	(0.677)	(0.927)	(0.645)	(0.995)
Sample Size	11,667		11,667		11,667		11,667	
Log Likelihood	-1,303.74		-2,514.31		-1,307.16		-1,305.09	
$\widehat{\Pr(y^* = 1)}$	0.223		0.255		0.207		0.207	
$\hat{a}_0$	0.514		0.570		0.470		0.483	
$\hat{a}_1$	0.019		0.017		0.018		0.018	

Quasi Maximum Likelihood Standard Errors in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively

Table 7: Robustness Checks - Property Crime

PROPERTY_B	No Weights (1)		No Exclusion (2)		Other Present (3)	
	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$
IMMIGRANT	-0.493**	-0.209	-0.232	-0.343	-0.143	-0.387
	(0.210)	(0.262)	(0.276)	(0.305)	(0.528)	(0.399)
AGE	-0.270***	-0.261***	-0.242***	-0.192*	-0.170	-0.244
	(0.035)	(0.066)	(0.083)	(0.107)	(0.285)	(0.367)
AGE^2	0.004***	0.004***	0.003***	0.003**	0.002	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)
MALE	0.464***	-0.179	0.401***	-0.239	0.382***	-0.168
	(0.125)	(0.130)	(0.149)	(0.169)	(0.122)	(0.394)
WHITE	0.487**	-0.074	0.369	-0.239	0.369*	-0.103
	(0.189)	(0.211)	(0.297)	(0.240)	(0.219)	(0.760)
SOUTH	0.050	0.151	0.175	0.200	0.085	0.188
	(0.162)	(0.178)	(0.247)	(0.194)	(0.456)	(0.558)
MIDLANDS	-0.032	0.222	0.066	0.126	0.022	0.086
	(0.177)	(0.184)	(0.203)	(0.208)	(0.212)	(0.355)
NORTH	0.103	0.299	0.379	0.438**	0.369	0.475*
	(0.201)	(0.192)	(0.297)	(0.181)	(0.219)	(0.266)
TRUTHFUL		1.181***				
		(0.197)				
OTHER_PRESENT						0.294
						(0.505)
CONSTANT	2.576***	2.509***	2.298	3.070***	0.985	2.906**
	(0.569)	(0.526)	(1.558)	(0.715)	(5.056)	(1.179)
Sample Size	11,658		11,658		11,658	
Log Likelihood	-2,441.32		-1,432.94		-1,430.670	
$\widehat{\Pr(y^* = 1)}$	0.260		0.291		0.170	
$\hat{a}_0$	0.650		0.684		0.475	
$\hat{a}_1$	0.015		0.005		0.002	

Quasi Maximum Likelihood Standard Errors in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively

Table 8: Interaction Terms

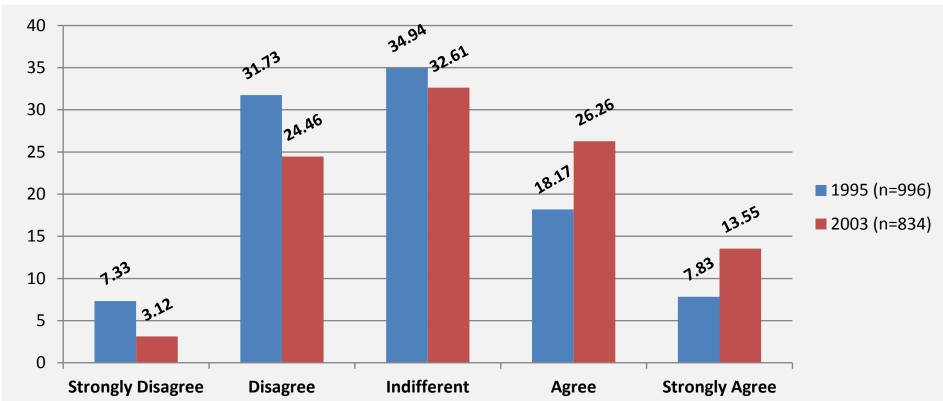
	By Region (1)		By Ethnic Group (2)	
	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$
Immigrant	-0.917***	-0.292	-0.196	-0.452
	(0.352)	(0.634)	(0.287)	(0.394)
Age	-0.248***	-0.258***	-0.252***	-0.192***
	(0.036)	(0.094)	(0.046)	(0.068)
Age^2	0.003***	0.004***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Male	0.521***	-0.192	0.428***	-0.265*
	(0.140)	(0.154)	(0.158)	(0.139)
White	0.571	-0.174		
	(0.372)	(0.392)		
Black			0.013	-0.067
			(0.367)	(0.438)
Asian&Other			-0.617	0.427
			(0.457)	(0.392)
Mixed			0.372	0.193
			(0.541)	(0.302)
South	0.009	0.175	0.275	0.264
	(0.217)	(0.258)	(0.237)	(0.227)
Midlands	-0.051	0.187	0.139	0.218
	(0.226)	(0.271)	(0.228)	(0.218)
North	0.301	0.547	0.489	0.488**
	(0.289)	(0.275)	(0.316)	(0.209)
Truthful		1.040***		0.822***
		(0.334)		(0.247)
Interaction Terms				
Immigrant*South	1.483***	0.531		
	(0.570)	(0.705)		
Immigrant*Midlands	0.690	0.199		
	(0.556)	(0.893)		
Immigrant*North	0.240	-2.079		
	(0.539)	(1.675)		
Immigrant*Black			-0.934	0.354
			(0.583)	(0.956)
Immigrant*Asian&Other			0.054	0.237
			(0.730)	(0.698)
Immigrant*Mixed			-0.615	0.323
			(0.919)	(0.836)
Constant	1.997**	2.578***	2.777***	2.070***
	(0.838)	(0.566)	(0.898)	(0.498)
Sample Size	11,658		11,658	
Log Likelihood	-1,413.96		-1,418.40	
$\widehat{\Pr(y^* = 1)}$	0.243		0.297	
$\hat{a}_0$	0.625		0.714	
$\hat{a}_1$	0.014		0.010	

Quasi Maximum Likelihood Standard Errors in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively.

Appendix: Attitudes Towards Immigrants in the UK

In this section we provide brief evidence on the attitudes of British citizens towards immigration and crime. For this purpose we utilise data from the 1995 and 2003 BSA cross-section surveys, where respondents indicated whether they agree or disagree with the statement that "immigrants increase crime rates" using a 5 points Likert-type scale (from 1="strongly agree" to 5= "strongly disagree"). Figure 1 very interestingly shows a clear shift of the observed unconditional probability distribution from 1995 to 2003 towards "agree/strongly agree that immigrants increase crime rates". More precisely, the percentage of people indicating that they agree or strongly agree jumped from 26% in 1995 to 40% in 2003. It is very interesting that, as shown in Figure 2, where we present the trends of total crime and immigrant population in the UK since 1990, this change in beliefs happened in a period where crime in the UK was falling, at least for total crime (see also, Chaplin, Flatley and Smith, 2011, where figures for recorded by police crime are also presented). However, at the same time it coincides with an increase in the percentage of immigrants' population from around 7% of the total population in 1993 to around 9% in 2003. This might point out that the continuous increase in immigrants' share is an important factor for this shift in the attitudes of British citizen towards immigration and crime.

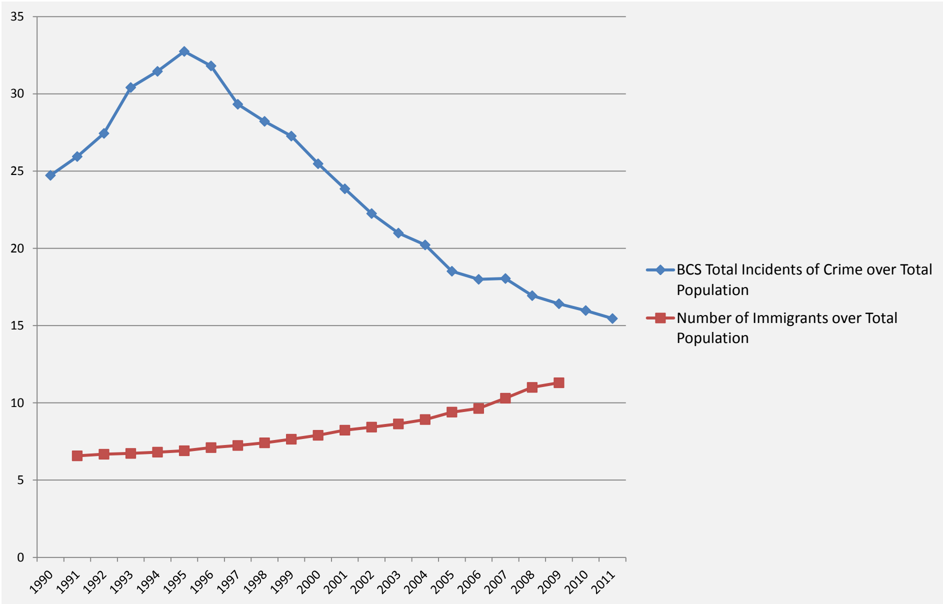
Figure 2: Do Immigrants Increase Crime Rates? UK data from the 1995 and 2003 BSA surveys



The results in Figure 1, however, do not recognise that there might be some differences between participants in 1995 and participants in 2003 also associated with attitudes to immigrants, such as differences in age, education and political ideology. This might reflect either differences just because of changes in the survey designs (the survey conductor changed from 1995 to 2003) or due to fundamental changes in Great Britain's population (for example, the population tends to become more educated and older). Therefore, we further perform a small regression analysis where we control for basic characteristics that might differ between respondents in 1995 and respondents in 2003 but also determine attitudes to immigration.[32] The results of an Ordinal Probit regression model are presented in the 3 columns of Table 1, where a simple dummy for year 2003 aims at capturing the evolution in respondents' attitudes. Note that the dependent

[32]Descriptive statistics of the variables used in these regressions are not presented here but are available on request.

Figure 3: Trends in Crime and Immigration*



* The BCS Total Incidents of Crime are taken from the Home Office Home Office Crime Statistics. Total Number of Immigrants and Total Population is takes from OECD.Stat Extracts.

variable is recoded such as value 1 now denotes "totally disagree ..." and value 5 denotes "totally agree ...". We include covariates for gender, age, education status, political ideology (specification 1), region, union, marital and employment status (specification 2), and citizenship status (specification 3).

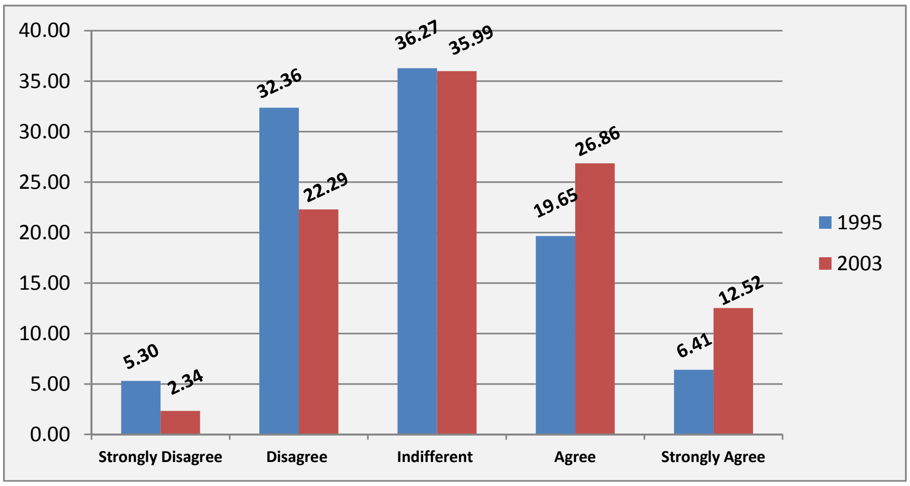
Although this small exercise produces some interesting results which would merit some discussion, here we concentrate on the effect of dummy YEAR_2003. From all 3 specification, it is clear that, holding the aforementioned observables constant, moving from 1995 to 2003 results in a strong increase in the sentiment that immigrants increase crime rates. To quantify the estimated effect of the dummy "Year 2003" we also calculated the predicted probabilities for the five categories (and the standard errors of these probabilities using the Delta method) conditional on the observed characteristics of specification 1, using a "representative" individual who is male, left-wing, and has got A-level or O-level qualification. These are presented in Figure 3 and indicate that moving from 1995 to 2003 the probability of responding with "disagree/strongly disagree that immigrants increased crime rates" decreased by 13 percentage points, but the probability of "agree/strongly agree ..." increased by 13.3 percentage points (these changes are significant at 1% significance level). Unfortunately, information on immigration status and ethnic group is not available in the data, but we expect that controlling for this would increase the estimated difference, as it is the case when controlling for citizenship status in specification (3). It is also interesting that even more recent data from the 2009 BSA show even stronger evidence of these negative beliefs, as around 81% of the respondents believe that "it is very likely, or somewhat likely, that more immigrants bring about higher crime rates" while only 19% believe that "it is not too likely or not likely at all" (these results are available on request).

Table 9: Ordinal Probit - Do Immigrants Increase Crime Rates?

	(1)	(2)	(3)
Year_2003	0.372***	0.380***	0.389***
	(0.052)	(0.053)	(0.054)
Age/100	0.827***	0.457*	0.418*
	(0.167)	(0.250)	(0.251)
Male	0.279***	0.312***	0.310***
	(0.053)	(0.060)	(0.060)
Education (No Secondary Qualification)			
CSE	-0.166*	-0.160	-0.160
	(0.100)	(0.101)	(0.103)
A-level,O-level	-0.306***	-0.295***	-0.309***
	(0.070)	(0.072)	(0.071)
Higher_Education_below_Degree	-0.525***	-0.492***	-0.495***
	(0.081)	(0.083)	(0.082)
Degree	-0.878***	-0.831***	-0.812***
	(0.094)	(0.098)	(0.098)
Political Ideology (Right-Wing)			
Left-Wing	-0.347***	-0.319***	-0.306***
	(0.064)	(0.066)	(0.066)
Centre	-0.275***	-0.274***	-0.278***
	(0.078)	(0.079)	(0.080)
Other_Party/Indifferent	-0.147*	-0.090	-0.080
	(0.085)	(0.089)	(0.089)
Dummies for Region, Union Membership, Marital Status, Working Status		✓	✓
Citizenship Status			
Only_Father_Citizen			0.606*
			(0.345)
Only_Mother_Citizen			0.544**
			(0.224)
Both_Parents_Citizens			0.705***
			(0.150)
Sample Size	1,748	1,732	1,720
Log Likelihood	-2,395.10	-2,358.06	-2,328.21

Quasi-maximum likelihood standard errors in parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively

Figure 4: Conditional Predicted Probabilities for 1995 and 2003*



* We control for Age, Gender, Education and Political Ideology.