ORIGINAL ARTICLE   **Open Access**

Immigration status and property crime: an application of estimators for underreported outcomes

Georgios Papadopoulos

Correspondence:
G.Papadopoulos@uea.ac.uk
School of Economics, University of
East Anglia, Norwich Research Park,
Norwich NR4 7TJ, UK

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. Models that account for underreporting are used, since this is a major concern in crime self-reports. The results indicate that the reported crime is substantially underreported, but if anything, immigrants are less likely to underreport than natives. More importantly, controlling for underreporting and basic demographics, the estimates across all model specifications, although imprecise, indicate that immigration status and property crime are negatively associated. We finally find that the estimated relationship between immigration status and property crime differs across regions and ethnic groups.

JEL Codes: K42, J15, J22, C25, C51

Keywords: Criminal Behaviour; Immigration; Self-reports; Underreporting; NB2-Logit

1   Introduction

International migration is a topic that has been heavily debated by policy makers, especially in countries that experienced important immigration inflows, such as the UK. Consequently, academic communities have devoted extensive research to understand the actual impact of immigration on several outcomes of both the host and home countries, including the effect of immigration on the labour market (Borjas 2003; Dustman et al. 2005; Card 2009) and the welfare state of the host countries (Borjas 1999), the impact of brain drain on the countries of origin (Beine et al. 2008) and the impact of ethnic diversity on economic performance (Alesina and La Ferrara 2005), to mention only a few. Following the substantial inflows and debates, the general public also started developing negative beliefs towards the migrant population, since they perceived immigrants not only as a major competitor in the labour market, but also as a main factor for deteriorating several problems of the host countries, especially crime. Indeed, at least for the UK, data from the British Social Attitudes survey (BSA), reveal that British citizens generally believe that immigrants increase crime rates (see, Appendix A 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. Interestingly, a high proportion of the empirical research does not share the hostile views dominating

individual opinions. For example, most research for the US indicates that if any, this association is negative (Butcher and Piehl 1998, 2007; Ousey and Hubrin 2009; Wadswarth 2010), while the results for Europe are mixed for property crime but no association is found for violent crime (Bianchi et al. 2012; Bell et al. 2013; Bell and Machin 2013; Jaitman and Machin 2013).

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]. Because underreporting of criminal activities is a major concern in crime self-reports, regression strategies that do not take into account this problem will result in inconsistent estimates of the determinants of criminal behaviour. Therefore, in this paper we tackle this problem by utilising regression models that attempt to control for underreporting. These models allow for consistent estimation of both the determinants of true criminal activity and the determinants of reporting behaviour by only using the information of observed self-reported crime.

Our estimates suggest that responses of criminal behaviour are considerably underreported. However, if anything, immigrants underreport by less than natives. In addition, once controlling for underreporting and basic demographic characteristics, we find that on average, immigrants are less involved in property crime, although the estimates are imprecise. Even though the estimated immigration-crime difference is not statistically significant in our baseline models, all sensitivity analysis shows that it is very robust. This 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. Finally, recognising the high heterogeneity of immigrant population, we investigate whether the immigration-crime association depends on certain groups of covariates, such as ethnic status or location. We actually find that immigrants who are located in London and black immigrants are significantly less criminally active than their native counterparts.

The remainder of this paper is organised as follows. Section 2 briefly discusses theoretical views on the immigration-crime link and presents a short literature review on the topic. Section 3 presents the regression models that allow for underreporting, while Section 4 discusses the data. Section 5 and 6 present the main results and the robustness checks, respectively. Section 7 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 8.

2   Theoretical views and a brief literature review on immigration and crime

According to economic theory there are two channels through which immigration and crime are linked. 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 UK, it has been found that there are minimal effects of immigration on the British labour market outcomes (see, Dustman et al. 2005; Manacorda et al. 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. The following discussion is about the latter, as this paper focuses on the individual level immigration-crime relationship.

The general view is that immigrants would find criminal activities more attractive, since their legal labour market opportunities are less favourable than natives' ones. For example, immigrants are on average poorer and more likely to be unemployed (see, for example, Algan et al. 2010). However, although economic models of crime agree that labour market outcomes are important in explaining participation in property crime, they also suggest that risk attitudes are equally important, as crime is a highly risky action that involves potential apprehension and punishment (for a simple economic model of criminal participation please refer to Appendix B).

In this direction, there are some reasons suggesting that immigrants may be less willing to take risks associated with criminal activities. For instance, holding everything else constant, crimes are more costly for immigrants, because a potential apprehension would jeopardise their smooth integration into the host country's society. It may even result in deportation, which according to Butcher and Piehl (2007) is an important disincentive to commit crimes. In addition, 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 immigrants may be more likely to be punished and face more severe punishments compared to natives, for the same crimes. Even if there is no discrimination, immigrants may be still more likely to be punished or receive more sever punishment, just because they lack knowledge regarding the UK criminal justice system, or because they have insufficient resources to acquire appropriate legal support. Finally, 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 (Sharp and Budd 2005), all else being equal.

Therefore, according to simple economic theory, the effect of immigration status on criminal activity can go either way. But what does the up-to-date empirical research indicate? As mentioned in the introduction, research for the US shows that in general immigration has a negative impact on crime, but evidence for Europe provides mixed results for property crime but suggests that no association exists for violent crime. Although there is some agreement on the statement above, the empirical evidence is still not robust, as different studies come to different conclusions, depending on the host countries being studied, the composition of immigrants, but most importantly, the statistical strategies and data each researcher uses to handle the research question.

Most researchers have used administrative crime panel data on an attempt to estimate the 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 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 staying in more expensive areas or because they tend to locate in areas where there is a large population of residents of the same ethnic background. Although panel data techniques alleviate this problem by controlling for unobserved time-invariant location characteristics (see, Butcher and Piehl 1998; Ousey and Hubrin 2009; Wadswarth 2010), they do not solve it completely, as there might always be unobserved time-varying features that affect both location decisions and crime rates. Some researchers attempted 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, Spenkuch 2013; Bianchi et al. 2012; Bell et al. 2013).

Of particular interest are the papers by Bell et al. (2013), Bell and Machin (2013) and Jaitman and Machin (2013), as to my best knowledge, these are the only papers that use data from the UK. Bell et al. (2013) examine the impact 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. However, they find that the A8 wave did not affect property crime. Bell and Machin (2013) study the effect of immigrant neighbourhood segregation on crime using mainly police recorded data. They find strong evidence across different specifications that, *ceteris paribus*, crime is lower in immigrant enclaves. In both papers, the findings suggest that there is no effect for violent crime. Finally, using spatial econometric strategies, Jaitman and Machin (2013) find no effect of immigration on crime both for A8 and non-A8 immigrants.

In a different direction, some researchers have focused on studying the individual-level relationship. Using official criminal records, such as arrests or prison records, they usually find that the proportion of immigrants in total imprisoned/arrested population is higher than the proportion of immigrants in the general population (see, Tonry 1997; Yeager 1997). 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 remains unrecorded, the so-called "dark figure" of crime (see, MacDonald 2002). If the recording mechanism is somehow biased against immigrants, then these statistics overestimate immigrants' involvement into criminal activities.

Another common practice among researchers to investigate the individual-level link has been the use of crime self-reports (see, for example, Junger-Tas and Marshall 1999). 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; Butcher and Piehl 1998). However, underreporting of responses is a major concern in these studies, since questions try to elicit information on a very sensitive part of personal behaviour. Therefore, if for some reasons immigrants tend to underreport 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 underreporting.

3   Regression models

As the observed information on property crimes is given in count form (see Section 4 for details), non-linear regression models for count data should be used. Conventional non-linear estimators, however, are inconsistent if underreporting, or more generally, response error in the outcome variable is present (see, Hausman et al. 1998; Cameron and Trivedi 1998; Winkelmann 2008). The problem is even more salient if underreporting depends on individual characteristics, which is certainly the case with crime self-reports. Therefore, consistent estimation of the parameters of interest requires the use of models that take into account underreporting. These are presented in the following subsection.

In addition, from the observed distribution of property crime, presented in Table 1, and its unconditional (weighted) mean and standard deviation presented in Table 2, we notice

Table 1 Distribution of property and violent crime

Num of crimes	Property		Violent	
	Count	Percent	Count	Percent
0	10,927	94.17	10,830	93.04
1	251	2.16	310	2.66
2	123	1.06	143	1.23
3	67	0.58	81	0.7
4	40	0.34	34	0.29
5	48	0.41	59	0.51
6–10	63	0.54	94	0.81
11–15	31	0.27	34	0.29
16–25	25	0.22	32	0.27
26–49	16	0.14	13	0.11
50–99	7	0.06	7	0.06
>100	6	0.05	3	0.03
Total	11,604	100	11,640	100

two important features that complicate estimation of count data models. Firstly, there is an exceptionally large concentration on outcome zero, as 94.17% of the respondents reported committing zero crimes. Secondly, positive outcomes are highly dispersed with a few extreme values. Because of this, we also provide some robustness checks utilising an estimator that uses only the binary choice information, whether or not someone has committed a crime and study whether immigrants are more or less likely than natives to commit property crimes once controlling for underreporting. The latter is presented in subsection 3.2.

3.1 The NB2-Logit and the ZI-NB2-Logit

Here we present 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 refer to Papadopoulos (2011).

To begin with, for individual i (with $i = 1, \ldots, n$), suppose that y_i^* is the number of committed crimes in a given period of time. We assume that y_i^*, conditional on a $(K \times 1)$ vector of covariates $\mathbf{x}_i$ follows the NB2 distribution with, $\lambda_i = E\left(y_i^*|\mathbf{x}_i\right) = \exp\left(\mathbf{x}_i'\boldsymbol{\delta}\right)$ and $\omega_i = Var\left(y_i^*|\mathbf{x}_i\right) = \lambda_i + \alpha\lambda_i^2$, where α captures gamma specific unobserved heterogeneity with $\alpha > 0$. We prefer the NB2 distribution to Poisson, since the equidispersion assumption ($\lambda_i = \omega_i$) 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 observed count of crime which are 0.34 and 34.32, respectively. We expect the conditional mean of true crime to also be substantially lower than the conditional variance.

However, since i might decide not to report a number of committed crimes, only a subset of y_i^* is observed. Now assume that the number of reported 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,

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

Table 2 Descriptive statistics (weighted means by immigrations status)

Variables	Definition	Weighted means			DIfference	Min	Max
		All	Natives	Immigrants			
Crime variables							
PROPERTY_c[a]	Reported number of committed property crimes in last year	0.342	0.366	0.160	-0.206***	0	225
		(5.858)	(6.011)	(2.218)			
PROPERTY_b	= 1 if respondent reported a property crime last year, = 0 otherwise	0.055	0.057	0.039	-0.018**	0	1
VIOLENT_c[b]	reported number of committed violent crimes in last year	0.286	0.310	0.107	-0.202***	0	512
		(4.132)	(4.252)	(0.918)			
VIOLENT_b	= 1 if respondent reported a violent crime last year, = 0 otherwise	0.054	0.056	0.036	-0.020***	0	1
Explanatory variables							
IMMIGRANT	= 1 if respondent lived all life in the UK, = 0 otherwise	0.119				0	1
AGE	Age of respondent in years	36.738	36.554	38.098	1.544***	10	66
MALE	= 1 if respondent is a male, = 0 otherwise	0.497	0.496	0.505	0.008	0	1
WHITE	= 1 if respondent is white, = 0 otherwise	0.909	0.956	0.553	-0.404***	0	1
BLACK	= 1 if respondent is black, = 0 otherwise	0.023	0.010	0.120	0.101***	0	1
MIXED	= 1 if respondent is of mixed ethnicity, = 0 otherwise	0.012	0.008	0.039	0.031***	0	1
ASIAN&OTHER[c]	= 1 if respondent is Asian or any other ethnicity, = 0 otherwise	0.056	0.022	0.214	0.262***	0	1
NORTH	= 1 if resp. lives in North or North West, or Yorks & Humperside, = 0 otherwise,	0.274	0.288	0.175	-0.112***	0	1
MIDLANDS	= 1 if resp. lives in East Midlands, or West Midlands, or Wales, = 0 otherwise	0.235	0.246	0.150	-0.097***	0	1
SOUTH	= 1 if resp. lives in East Anglia, or South East, or South West, = 0 otherwise	0.351	0.358	0.298	-0.060***	0	1
LONDON[d]	= 1 if respondent lives in London, = 0 otherwise	0.139	0.107	0.376	0.268***	0	1
DEPRIVATION[e]	Integer values from 1 [= least deprived] to 10 [= most deprived]	5.500	5.426	6.023	0.597***	1	10
TRUTHFUL	= 1 if respondent said that answered crime questions truthfully, = 0 otherwise	0.942	0.946	0.915	-0.031***	0	1
OTHER_PRESENT	= 1 if someone else present during self-completions, = 0 otherwise	0.285	0.288	0.263	-0.025*	0	1

[a]Property crime consists of: thefts and attempted vehicle thefts, thefts and attempted thefts of parts from inside or outside vehicles, robberies, burglaries, thefts form person, work, school, shops, other thefts and criminal damage.
[b]Violent crime consists of: assaults with and without injuries and robberies. Note that only 9 individuals reported committing a robbery.
[c]50% of respondents that belong to OTHER ethnic group are Chinese. This is why people who belong to ASIAN or OTHER ethnicity are grouped in one variable.
[d]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.
[e]This is the Multiple Deprivation Index of England, constructed as a weighted mixture of the individual deprivation indices of income, employment, health, education, crime, barriers to housing and living environment at the Lower Super Output Areas (LSOAs). Very briefly, this index takes integer values from 1 [least deprived] to 10 [most deprived] indicating the decile of deprivation in which a respondent scores. For example, if someone scores at the 7th decile, only 30% of the population resides in more deprived areas. Each respondent, depending on the LSOA he/she resides, is matched by the Home Office to the corresponding decile of this variable.
Weighted Standard Deviations in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively.

Putting this in a regression framework, we assume that the probability to report a committed crime, p_i, follows the Logit model with $p_i = \Pr(c_{ij} = 1|\mathbf{z}_i) = \Lambda\left(\mathbf{z}_i'\boldsymbol{\eta}\right) = \exp\left(\mathbf{z}_i'\boldsymbol{\eta}\right) / \left[1 + \exp\left(\mathbf{z}_i'\boldsymbol{\eta}\right)\right]$, where $\mathbf{z}_i$ is another $(L \times 1)$ vector of covariates.

If we further assume that y_i^* and c_{ij} are conditionally independent, we can show, for example 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,

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

respectively (see, Papadopoulos 2011). This is the NB2-Logit model, with PDF given by,

$$\Pr\left(Y_i = y_i|\mathbf{x}_i, \mathbf{z}_i\right) = \frac{\Gamma\left(y_i + \alpha^{-1}\right)}{\Gamma\left(y_i + 1\right)\Gamma\left(\alpha^{-1}\right)}\left(1 + \alpha\,\mu_i\right)^{-\left(\alpha^{-1}+y_i\right)}\left(\alpha\,\mu_i\right)^{y_i}. \tag{3}$$

We can estimate $\boldsymbol{\theta} = (\boldsymbol{\delta}, \boldsymbol{\eta}, \alpha)$ using the Maximum Likelihood Estimator (MLE) since we can easily obtain the the log-likelihood function as,

$$\ln\mathscr{L}(\boldsymbol{\theta}) = \sum_{i=1}^{n}\left(\ln\left(\frac{\Gamma\left(y_i + \alpha^{-1}\right)}{\Gamma\left(y_i + 1\right)\Gamma\left(\alpha^{-1}\right)}\right) - \left(\alpha^{-1} + y_i\right)\ln\left(1 + \alpha\,\mu_i\right) + y_i\left(\ln\mu_i + \ln\alpha\right)\right). \tag{4}$$

Maximization of (4), using optimisation algorithms such as the Newton-Raphson, yields consistent estimates of $\boldsymbol{\theta}$, given correct specification of the model; that is, the true data generating process (DGP) is NB2-Logit. Alternatively, if we assumed that y_i^* followed the Poisson distribution, we would end up with a Poisson-Logit model, which belongs to the Linear Exponential Family (LEF) (see Gourieroux et al. 1984; Staub and Winkelmann 2013), and only correct specification of the first moment would be enough for consistency. On the other hand, the NB2-Logit belongs to the LEF only if α is known. Since α is subject to estimation, NB2-Logit is not an LEF and therefore, misspecification of higher moments than the mean leads to inconsistency. However, under the presence of high overdispersion, as expected here, the NB2-Logit is highly more efficient than the Poisson-Logit.

Note that, depending on the application under investigation, $\mathbf{x}_i$ and $\mathbf{z}_i$ may be identical, overlapping or disjoint. As Papadopoulos and Santos Silva (2008) show, however, unless appropriate restrictions are imposed, identification of $\boldsymbol{\theta}$ 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 exclusion restriction on the crime process (NB2 part), meaning that there is one variable which, conditional on the covariates, has a significant impact on the reporting process but no impact on the crime process. We will call this, a *strong* restriction. 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 dubious[3].

We need to stress that the estimated parameters of the conditional expectation of this model are observationally equivalent to the estimates of a Zero-Inflation NB2 without underreporting, where the probability of inflation is given by $\Lambda\left(\mathbf{z}_i'(\boldsymbol{\eta})\right)$. According to this framework there is a proportion of individuals who, regardless of their characteristics,

never participate in criminal activities and therefore do not commit any crimes (*structural zeroes*), while the rest of them may participate in crime but still, they may commit zero crimes (*incidental zeroes*) or a positive number of crimes. Then, an estimated coefficient of $\boldsymbol{\eta}$ could be interpreted in two different ways. For instance, if we are interested in the effect of being a male and estimate a positive coefficient on the reporting process, this could mean that males are more likely than females to report a committed crime, or less likely to never participate in crime (that is, less likely to belong to the zero-inflation (ZI) category).

Nevertheless, NB2-Logit can be extended to a model that disentangles ZI from under-reporting. According to this, let ξ be the probability of being an individual that does not participate in crime. Therefore, there is probability $(1 - \xi)$ to be an individual who may commit crimes, but his/her responses are also subject to underreporting; that is, they follow the NB2-Logit model. In our regression framework, the probability of ZI, conditional on a set of characteristics $\mathbf{q}_i$, also follows a Logit model with $\xi_i = \exp\left(\mathbf{q}_i'\boldsymbol{\zeta}\right) / \left(1 + \exp\left(\mathbf{q}_i'\boldsymbol{\zeta}\right)\right)$. Therefore, the conditional probabilities of zero and positive outcomes are given by,

$$
\begin{aligned}
&\Pr(y_i = 0|\mathbf{x}_i, \mathbf{z}_i) = \xi_i + (1 - \xi_i)\left(1 + \alpha\,\mu_i\right)^{\alpha^{-1}}, \\
&\Pr(y_i > 0|\mathbf{x}_i, \mathbf{z}_i) = (1 - \xi_i)\frac{\Gamma\left(y_i + \alpha^{-1}\right)}{\Gamma\left(y_i + 1\right)\Gamma\left(\alpha^{-1}\right)}\left(1 + \alpha\,\mu_i\right)^{-\left(\alpha^{-1}+y_i\right)}\left(\alpha\,\mu_i\right)^{y_i},
\end{aligned}
\tag{5}
$$

respectively, where $(1 + \alpha\,\mu_i)^{\alpha^{-1}}$ is the probability of a zero outcome from the NB2-Logit model. The log-likelihood function is given by,

$$
\begin{aligned}
\ln\mathscr{L}(\boldsymbol{\theta},\boldsymbol{\zeta}) = &-\sum_{i=1}^{n}\ln\left(1 + \exp\left(\mathbf{q}_i'\boldsymbol{\zeta}\right)\right) + \sum_{y=0}\ln\left(\exp\left(\mathbf{q}_i'\boldsymbol{\zeta}\right) + \left(1 + \alpha\,\mu_i\right)^{\alpha^{-1}}\right) \\
&+ \sum_{y>0}\ln\left(\frac{\Gamma\left(y_i + \alpha^{-1}\right)}{\Gamma\left(y_i + 1\right)\Gamma\left(\alpha^{-1}\right)}\right) - \left(\alpha^{-1} + y_i\right)\ln\left(1 + \alpha\,\mu_i\right) + y_i\left(\ln\mu_i + \ln\alpha\right).
\end{aligned}
\tag{6}
$$

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

3.2 The misclassification Probit model

The Misclassification Probit model (MisProbit) presented here is based on the model developed by Hausman et al. (1998) and the reader may refer to it for details. It arises naturally from a latent variable specification. Assume that 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 $\mathbf{x}_i$ such that,

$$
U_i^* = \mathbf{x}_i'\boldsymbol{\beta} + \epsilon_i, \quad \text{for, } i = 1,\dots n, \quad \text{and, } \epsilon_i|\mathbf{x}_i \sim \mathcal{N}(0, 1).
\tag{7}
$$

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

$$
y_{b,i}^* = 1 \text{ if } U_i^* > 0, \quad \text{and } y_{b,i}^* = 0 \text{ if } U_i^* \leq 0,
\tag{8}
$$

where y_b^* 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,

$$
\Pr\left(y_{b,i}^* = 1|\mathbf{x}_i\right) = \Pr\left(U_i^* > 0|\mathbf{x}_i\right) = \Pr\left(\epsilon_i > -\mathbf{x}_i'\boldsymbol{\beta}|\mathbf{x}_i\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right),
\tag{9}
$$

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

Suppose now that the reported crime, denoted by the binary variable $y_{b,i}$, does not coincide with $y_{b,i}^*$ since there is a probability of underreporting (in this context is referred to as misclassification of a true one as a zero)[4]. We can therefore write the probability of reporting at least one committed crime as, $\Pr\left(y_{b,i} = 1 | y_{b,i}^* = 1, \mathbf{z}_i\right) = \Lambda\left(\mathbf{z}_i'\boldsymbol{\gamma}\right)$, where we assumed that this also depends on $\mathbf{z}_i$ according to a Logit model. Therefore, the conditional probability of observing a crime is given by,

$$\Pr\left(y_i = 1 | \mathbf{x}_i, \mathbf{z}_i\right) = \Lambda\left(\mathbf{z}_i'\boldsymbol{\gamma}\right)\Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right). \tag{10}$$

From (10), the log-likelihood function is obtained as,

$$\ln\mathscr{L}(\boldsymbol{\vartheta}) = \sum_{i=1}^{n}\left(y_i\ln\left[\Pr\left(y_i = 1 | \mathbf{x}_i, \mathbf{z}_i\right)\right] + (1 - y_i)\ln\left[1 - \Pr\left(y_i = 1 | \mathbf{x}_i, \mathbf{z}_i\right)\right]\right). \tag{11}$$

Given that the probability model above describes the true DGP, maximization of (11) consistently estimates both $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the determinants of the probability to commit a crime and the probability to correctly report it, respectively[5].

We need to stress that the functional form of this model allows identification even without exclusion restrictions, although exclusion restrictions could always facilitate the estimation procedure, particularly in applications where there is not a lot of variation in the dependent variable. Note also that as was the case with the count data models, exactly the same model can be obtained under a ZI framework, with probability of ZI equal to $\Lambda\left(\mathbf{z}_i'(-\boldsymbol{\gamma})\right)$. Thus, we need to be careful when interpreting the estimated coefficients of this model.

4 The CJS data and discussion of variables

As mentioned in the introduction, for our analysis we use the Crime and Justice Survey (CJS) of 2003, a national representative survey where respondents in England and Wales were asked questions regarding their criminal activities. This survey uses computer-based self-completions when it comes to questions related to crime, as opposed to face-to-face interviews, a method proven to increase reliability of responses (see, Turner et al. 1998). However, respondents may have still been reluctant to reveal information about their criminal behaviour and therefore, some degree of underreporting in the data is still expected. To have a rough idea about the level of underreporting 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 et al. (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 UK population surveys, such as the LFS 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 are likely to be more prone to crime than participants. However, to correct for potential sample selection biases requires having rich information on non-respondents, which is not available. Therefore, this problem is ignored in the analysis, hoping that conditional on the covariates, selection becomes random[6].

The main outcome variable of our analysis, PROPERTY_C, is the reported number of committed property crimes during the twelve months prior to the interview. We also have information on the count of violent crimes, VIOLENT_C, for the same reference period, which is used to check the robustness of our main results. The distribution of the crime variables is presented in Table 1, while definitions together with descriptive statistics are provided in Table 2.

At this point, it is worth noting that some effort is made to keep the sample size as large as possible, because, as mentioned already, not only are the econometric methods used in the empirical analysis quite demanding, but also 94.17% of respondents reported no property crime with the remaining positives being very dispersed. Therefore, larger samples assist in estimating the coefficients of interest more precisely. For this purpose, apart from the core sample (representative sample of people between 10 and 65 years old), we also use 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. A tabulation by sample type is given in Table 3. Our final final sample consists of all individuals between 10 and 65 years of age, for whom we have information on the count of property crime. So, we end up with 11,604 individuals, 5,570 males and 6,034 females. Note that the sample size differs between the count and the binary form of the property crime variable because some respondents who reported a crime did not report the number of crimes they committed.

Main interest of this study lies in estimating the differences in criminal activity and reporting behaviour between immigrants and natives. While it is common in empirical studies to define an immigrant as a person who is born outside the reference country, country of birth is not available in the questionnaire. Instead, to construct the dummy variable IMMIGRANT we used the following: "Can I just check how long have you lived in the United Kingdom?". Respondents that replied "All my life", are considered natives; otherwise, they are classified as immigrants[7]. Although we only have 728 immigrants in the core sample, their number increases to 2,006 by exploiting the *youth-boost* and, most importantly, *ethnic-boost* samples.

Although the CJS provides a rich set of respondents' characteristics, such as employment status, education, parental characteristics and perceived risks, we only use controls for basic demographics. Thus, apart from IMMIGRANT, the following explanatory variables are used: AGE; MALE; five ethnic group dummies: WHITE, BLACK, ASIAN&OTHER and MIXED; four regional dummies: NORTH, MIDLANDS, SOUTH, LONDON; and the Multiple Index of Deprivation, DEPRIVATION. Definitions together with explanations and descriptive statistics are given in Table 2. Note that, unless there are very good

Table 3 Tabulation of CJS respondents by sample type

Sample type	Total		Immigrants		Natives	
	N	%	N	%	N	%
Core 10-65	6,753	58.20	728	36.29	6,025	62.77
Youth boost 10-25	3,069	26.45	185	9.22	2,884	30.05
Ethnic boost 10-65	1,782	15.36	1,093	54.48	689	7.18
Total	11,604	100	2,006	100	9,598	100

arguments for imposing exclusion restrictions, logically both the crime and the reporting processes are functions of the same variables. For instance, age influences both crime and reporting behaviour, as younger people commit on average more crimes and might also be less willing to reveal their true criminal behaviour, for example because of parental pressure. Similar arguments can be made for the other independent variables.

Of course, it would be interesting to explore the behaviour of the estimated immigration-crime differentials once controlling for labour market outcomes, risk attitudes, etc. However, this information is not used in the empirical analysis because of two reasons. Firstly, most of these variables are derived from questions replied only by people older than 17 years old, which results in reducing the sample by around 2,500 individuals and increasing the percentage of zeroes to 95.54%. 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 underreporting are used, the variation of the reduced sample is not enough to allow identification of all parameters of interest[8]. Secondly, all of these variables 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 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 significant impact 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 differentials between immigrants and natives, once controlling for underreporting and the basic demographics listed above[9].

As described in Section 3.1, one strategy for identification of 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 two variables that can be used for this purpose.

Firstly, respondents were asked whether they replied to the questions related to crime truthfully, which is used to create the dummy variable TRUTHFUL. 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 could not have affected criminal activity prior to the survey. If any empirical relationship exists, it would be because perceived truthfulness is correlated with unobserved characteristics that affect criminal behaviour, or because there is reverse causality of committed crimes on truthfulness[10]. 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.

Alternatively, in 32% of the interviews (3,768 observations) there was someone else present during the interview, mostly in the cases of individuals younger than 17 years old.

There is evidence, at least for face-to-face interviews (see, Aquilino 1993) that someone else's presence during responding to sensitive questions affects the reporting behaviour. Therefore, we created the dummy OTHER_PRESENT which can be arguably assumed to only affect the reporting behaviour. 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 it was also stressed by the interviewers that nobody should disturb the interviewee during the self-completion part.

Therefore, the baseline results are obtained exploiting the variable TRUTHFUL only, but results using the OTHER_PRESENT variable, or even no exclusion restrictions, are also presented in the robustness analysis section.

5   Main results

To start with, Table 2 shows that without controlling for demographic differences, the average number of crimes immigrants reported is 0.160, while this number is 0.366 for natives, a difference of 0.206 crimes that is significant at 1% significance level. Moreover, immigrants are 1.8 percentage points less likely to report a crime, which is significant at 5% level. So, immigrants are considerably less likely to report crimes. However, the results in specification (1) of Table 4, which presents the conventional NB2 estimates ($\tilde{\delta}$), and specification (1) of Table 5, which presents the conventional Probit ones ($\tilde{\beta}$), show that once we control for basic demographics, immigration status estimated coefficient becomes insignificant although it retains the negative sign. Of course, these results are not very informative, since the estimated differences may reflect either a true difference in criminal activities, or merely differences in reporting behaviour.

For this reason, we now turn our attention to the models that control for underreporting. Specifications (2), (3) and (4) of Table 4 present the results of NB2-Logit, where $\hat{\delta}$ and $\hat{\eta}$ are the estimates corresponding to criminal and reporting behaviour respectively, while specification (5) presents the ZI-NB2-Logit, where $\hat{\tilde{\zeta}}$ gives the estimates for ZI.

Concerning the main objective of this paper, the results from specification (2) show that after controlling for potential differences in the reporting behaviour between immigrants and natives, the effect of immigration status on properly crime remains negative, becomes larger in magnitude, but is still insignificant (it is significant only at around 30% level of significance). The increase in the magnitude of the estimated difference may be attributed to the fact that immigration status exhibits a positive coefficient on the reporting process, although insignificant as well, indicating that native-born individuals may underreport by more than immigrants. Specification (3) shows that after controlling for ethnicity, the effect of immigration status becomes only slightly smaller, as being white only has a slight positive but insignificant impact on the crime process[11]. In specification (4) we also add the overall index of multiple deprivation to control for other geographical differences associated with criminal behaviour. Because this index was not available for Wales in 2003, we had to drop all respondents residing in Wales, reducing the sample by 664 individuals. We find however, that this index has no significant effect on either criminal or reporting behaviour[12]. We therefore keep specification (3) as our *baseline* model, which is used for all robustness checks as well, as we would not like to drop all individuals from Wales.

Table 4 Main results - negative binomial 2 models

PROPERTY_C	NB2 (1)	NB2-Logit (2)		NB2-Logit (3)		NB2-Logit (4)		ZI-NB2-Logit (5)		
	$\tilde{\delta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\xi}$
IMMIGRANT	-0.440	-0.697	0.484	-0.617	0.451	-0.500	0.325	-0.745	0.109	-0.515
	(0.356)	(0.611)	(1.057)	(0.636)	(1.043)	(0.591)	(0.924)	(0.696)	(0.832)	(0.548)
AGE	0.0005	-0.662***	1.007***	-0.677***	1.008***	-0.624*	0.980***	-0.683**	1.045***	0.123**
	(0.031)	(0.198)	(0.197)	(0.224)	(0.206)	(0.325)	(0.302)	(0.298)	(0.219)	(0.058)
AGE^2	-0.001*	0.009***	-0.015***	0.009***	-0.015***	0.008*	-0.014***	0.010**	-0.015***	-0.001
	(0.000)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.003)	(0.001)
MALE	0.669**	1.428***	-1.317*	1.457***	-1.333*	1.234***	-0.906*	0.634	-0.404	-0.960***
	(0.275)	(0.430)	(0.685)	(0.443)	(0.687)	(0.464)	(0.825)	(0.470)	(0.604)	(0.259)
WHITE				0.196	0.638	0.276	0.638	-0.870	0.940	-1.605***
				(0.611)	(1.062)	(0.614)	(1.155)	(0.682)	(0.985)	(0.375)
SOUTH	0.504*	1.077*	-1.419*	1.024	-1.508	0.966	-1.600**	1.510**	-1.875**	0.516
	(0.284)	(0.601)	(0.814)	(0.696)	(0.937)	(0.599)	(0.786)	(0.680)	(0.759)	(0.474)
MIDLANDS	0.593**	0.272	0.418	0.247	0.255	0.434	0.273	0.553	-0.500	0.217
	(0.272)	(0.525)	(0.880)	(0.561)	(0.954)	(0.563)	(1.105)	(0.657)	(0.857)	(0.546)
NORTH	1.256**	1.717***	-1.527	1.676**	-1.628	1.943**	-2.266	2.715***	-2.776***	0.843*
	(0.012)	(0.594)	(0.956)	(0.662)	(1.027)	(0.864)	(1.435)	(0.894)	(0.964)	(0.470)
DEPRIVATION						0.053	-0.175			
						(0.096)	(0.163)			
TRUTHFUL			-1.282***		-1.237***		-1.302***		-1.757***	0.441
			(0.482)		(0.475)		(0.437)		(0.533)	(0.373)
CONSTANT	-1.299***	8.534***	-11.648***	8.633**	-12.253***	7.555	-10.422*	10.499**	-13.083***	-1.315
	(0.505)	(3.028)	(3.008)	(3.671)	(3.510)	(5.518)	(5.872)	(5.081)	(4.294)	(1.014)
Sample size	11,604	11,604		11,604		10,940		11,604		
Log likelihood	-2,344.45	-2,315.13		-2,313.99		-2,197.14		-2257.79		
$\hat{\alpha}$	46.39***	41.66***		41.64***		41.82***		14.589***		
$\hat{p}$		0.447		0.430		0.441		0.391		
$\hat{\xi}$								0.636		

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

Table 5 Robustness checks - binary models

	Probit					MisProbit				
		No ethnicity		Baseline	Depr.Index	No Excl.	Other present	No weights	10-25	Violent
	(1)	(2)		(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$\tilde{\beta}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
IMMIGRANT	-0.127	-0.429*	0.598	-0.309	-0.311	-0.351	-0.246	-0.434***	-0.243	-0.147
	(0.102)	(0.225)	(0.485)	(0.235)	(0.232)	(0.271)	(0.393)	(0.159)	(0.245)	(0.277)
AGE	-.0178**	-0.264***	0.323***	-0.268***	-0.261***	-0.288***	-0.226	-0.279***	-0.194***	-0.332***
	(0.007)	(0.049)	(0.075)	(0.049)	(0.046)	(0.044)	(0.158)	(0.037)	(0.031)	(0.042)
AGE^2	0.0001	0.004***	-0.005***	0.003***	0.004***	0.004***	0.003	0.004***		0.005***
	(0.0001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)		(0.001)
MALE	0.362***	0.318**	0.458**	0.321**	0.277*	0.308**	0.345*	0.335***	0.084	0.149
	(0.049)	(0.139)	(0.227)	(0.140)	(0.147)	(0.152)	(0.201)	(0.112)	(0.197)	(0.133)
WHITE				0.245	0.183	0.176	0.284	0.249	0.448*	0.259
				(0.201)	(0.199)	(0.248)	(0.290)	(0.158)	(0.259)	(0.197)
SOUTH	0.089	0.308	-0.394	0.27	0.256	0.272	0.185	0.104	-0.082	-0.282
	(0.082)	(0.198)	(0.330)	(0.203)	(0.196)	(0.238)	(0.373)	(0.144)	(0.259)	(0.276)
MIDLANDS	0.061	0.14	-0.208	0.124	0.225	0.098	0.056	0.082	-0.07	-0.184
	(0.084)	(0.188)	(0.340)	(0.200)	(0.221)	(0.214)	(0.226)	(0.171)	(0.268)	(0.274)
NORTH	0.066	0.531*	-0.796**	0.509**	0.543**	0.526*	0.362	0.21	-0.021	-0.197
	(0.086)	(0.243)	(0.336)	(0.260)	(0.248)	(0.311)	(0.571)	(0.171)	(0.28)	(0.282)
DEPRIVATION					-0.033					
					(0.028)					
CONSTANT	-1.204***	2.597***	-3.742***	2.867***	3.046***	3.328***	1.999	3.098***	2.979***	4.181***
	(0.127)	(0.895)	(0.608)	(0.935)	(0.896)	(0.928)	(3.150)	(0.623)	(0.890)	(0.789)
TRUTHFUL			-1.211***	-1.231***	-1.280***			-1.573***	-1.739***	-1.213***
			(0.267)	(0.264)	(0.250)			(0.202)	(0.287)	(0.276)
OTHER PRESENT							-0.393			
							(0.368)			
Sample size	11,658	11,658		11,658	10,993	11,658	11,658	11,658	5,023	11,667
Log likelihood	-1,452.88	-1,427.97		-1,422.51	-1,343.88	-1,432.72	-1,430.48	-2,444.18	-650.08	-1,310.73
Prob of committing a crime	0.064	0.370		0.370	0.370	0.422	0.279	0.362	0.514	0.368
Prob of reporting a crime		0.290		0.272	0.270	0.224	0.365	0.289	0.298	0.300

Note: because of space limitations, only specification (2) presents both the estimates from the crime equation ($\hat{\beta}$) and reporting equation ($\hat{\gamma}$). Specifications (3) to (9), although also obtained using the MisProbit model, only present the estimates for criminal behaviour. The presented estimated effects of TRUTHFULNESS and OTHER PRESENT are however from the reporting equation.
Quasi Maximum Likelihood Standard Errors in Parentheses.
*, **, *** denote 10%, 5%, 1% level of significance respectively.

We noted above that immigrants may be more likely to report a committed crime than a native. However, since the coefficients of the reporting process can also take a ZI interpretation, this positive coefficient might also mean that immigrants are less likely to belong to the group of people that never participate in crime. The ZI-NB2-Logit model in specification (5) resolves this issue as it disentangles underreporting from ZI. This model indicates that, indeed, immigrants are less likely to be in the ZI category (although this difference is again insignificant) and the effect of immigration status on reporting a committed crime drops both in magnitude and significance (even though it retains its sign). In addition, after controlling for both ZI and underreporting, the effect of immigration status becomes even larger in magnitude, but still statistically insignificant as the precision of the estimates decreases.

To quantify the above estimated differences, using specification (3), we calculated the predicted expected number of committed crimes both for a native and an immigrant, holding the other characteristics fixed according to a *representative* individual who is 25 years old, male, white and lives in London. We find that the expected number of crimes for the *representative* native is predicted to be 0.458 while this number is 0.247 for the *representative* immigrant. So the difference is around 0.211 crimes, while this difference for the conventional NB2 model was 0.112 crimes. So, although statistically insignificant, these differences are relatively large in terms of magnitude[13].

At this stage, it is worth mentioning some other interesting features from our results. First of all, we notice the large value of $\hat{\alpha}$, which is statistically significant at any significance level. Therefore, there is evidence that the data are highly over-dispersed even after conditioning on the regressors. As far as the reporting process is concerned, this model predicts that the average conditional probability of reporting a committed crime, calculated as $\hat{p} = \sum_{i=1}^{n} \Lambda \left(\mathbf{z}_i' \hat{\eta} \right) / n$, is around 44%. However, remember that $1 - \hat{p}$ can be interpreted as the proportion of individuals who do not participate in crime (ZI), which is 56%. The ZI-NB2-Logit sorts out this issue. Since it allows for ZI, the Logit process measures the probability of reporting a committed crime only for those who may choose to commit crimes. We interestingly find that the predicted average probability of ZI, calculated as $\hat{\hat{\xi}} = \sum_{i=1}^{n} \Lambda \left(\mathbf{q}_i' \hat{\hat{\xi}} \right) / n$, is around 64% (which is close to $1 - \hat{p}$) and that for those who may participate in criminal activities, the probability to report a committed crime is around 39%, which is similar to the figure suggested by NB2-Logit.

As discussed in the previous section, an exclusion restriction from the Logit process is required for identification of the NB2-Logit and ZI-NB2-Logit models. We therefore used the variable TRUTHFUL, which has a negative and strongly significant effect. Although these models are globally identified since TRUTHFUL is a *strong* exclusion restriction, we must still be very cautious since more than one maxima may exist. Indeed, regarding specifications (3) and (5), our investigations showed that two local maxima exist with log likelihood values equal to -2,315.80 and -2,261.05 respectively. These local maxima correspond to estimates that are very different from the ones of the global maxima shown in Table 4 (see subsection 6.2 for an example). Now, the negative effect of TRUTHFUL indicates that either, *truthful* respondents underreport by more than *non-truthful* ones, or that, in fact, *truthful* respondents are more likely to never participate in crime. However, the ZI-NB2-Logit results suggest that actually the former is more likely, as the impact of TRUTHFUL on the ZI process is insignificant, even though positive.

It would be also interesting to briefly discuss the effects of the other explanatory variables. To begin with, AGE seems to have a quadratic significant effect on both crime and reporting processes, where crime decreases with age in a decreasing rate, but the probability to report a committed crime increases in a decreasing rate. A quick calculation shows that crime reaches a minimum at about 36 years of age, while probability of reporting reaches a maximum at around 34 years of age. Note that we experimented with models that use higher polynomial of age, but none of them provided a better fit. Regarding the gender effect, the NB2-Logit suggests that being a male increases the average number of crimes, as someone would expect. The effect of being male on the reporting process is negative, which is puzzling, since this indicates that males are less willing to report their criminal activities, or that they are more likely to be in the ZI category. Nevertheless, the estimates of ZI-NB2-Logit reveal the opposite; that is, males are actually less likely to be in the ZI category and insignificantly less likely to report their crimes. Finally, the results of the regional dummies suggest that people who live in South or North commit more property crime and are less willing to report crimes compared to people who live in London.

6 Robustness checks

6.1 Results of MisProbit

We now turn to our MisProbit results, where we examine the effect of the immigration status dummy on the probability to commit a property crime, once we control for misclassification. The results are presented in Table 5, where $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ give the estimates of the crime process and the reporting process, respectively. In this section we discuss specifications (2) to (4) which correspond to specifications (2) to (5) from Table 4.

Overall, the MisProbit estimates back up the results of our count data models. We can see that the immigration status coefficient is again negative and actually statistically significant at 10% in specification (2) (p-value is 0.056). However, after controlling for ethnicity, although still negative and fairly large, it becomes insignificant (p-value increases to 0.188). This is because, the effect of WHITE on crime is positive (though insignificant), while immigrants are more likely to be non-white. Thus, once again, although there seems to be a negative relationship between immigration status and criminal behaviour, this is estimated imprecisely. We moreover see that immigrants are (insignificantly) more likely to correctly report a committed crime, or that, according to a ZI framework, immigrants are less likely to be in the ZI group. To quantify the above estimates, we calculate the predicted probabilities of committing a crime for our *representative* individual. Before controlling for ethnicity, this figure is 0.15 for a native and 0.07 for an immigrant, so that natives are twice as likely to commit a property crime. After controlling for ethnicity, these figures become 0.16 and 0.10 respectively.

Regarding other features of the MisProbit model, it predicts that the probability of reporting one of the committed crimes is around 30%, or that the probability of ZI is 70%. So, according to this model, people underreport more than what the count data models suggest, or they are more likely to be in the ZI category. Finally, the MisProbit model predicts that the average probability of committing a property crime, calculated as $\Pr \widehat{(y_i^* = 1)} = \sum_{i=1}^{n} \Phi \left(\mathbf{x}_i' \hat{\boldsymbol{\beta}} \right) /n$, is around 37%, which is much higher than the predicted average probability of the simple Probit model, calculated to be only 6.4%. However, interpreting MisProbit as a ZI model, this is actually the predicted

probability of committing a crime only for those that may participate in criminal activities.

6.2 Are the results driven by the exclusion restriction?

In this subsection I briefly intend to explain why our main results are not driven by the exclusion of variable TRUTHFUL from the crime process. Regarding our count data models, specifications (1) and (2) of Table 6 present results of including the dummy OTHER PRESENT in the reporting process of NB2-Logit instead of TRUTHFUL. Results of ZI-NB2-Logit, which are available on request, are very similar. As can be seen from (1), OTHER PRESENT has an insignificant effect on the probability to report a committed property crime. Therefore, we see from (2) that another maximum exists, which is very close, in terms of the log likelihood value, to this (global) maximum. As it is clear from (2), the second maximum corresponds to very different parameter estimates. As Papadopoulos and Santos Silva (2008) show, it appears that there is a close relationship between the parameters of the two maxima. Given that $\theta = (\delta, \eta)$ is the set of true parameters of the model, if the exclusion restriction is not *strong*, another maximum very close to the true one exists with parameter values $\tilde{\theta} \simeq (\delta + \eta, -\eta)$. The stronger the exclusion (for example, the case of *truthfulness*), the easier it is to distinguish the correct maximum based on the log-likelihood values, and the higher the deviation of $\tilde{\theta}$ from $(\delta + \eta, -\eta)$. Despite the fact that the likelihood values of (1) and (2) are too close, if we accept that (1) gives the correct maximum, the estimated parameters are very similar to the ones of our *baseline* model.

Now, regarding our binary models, specification (5) of Table 5 shows that the MisProbit model produces estimates very similar to the baseline model of specification (3) even without any exclusion restrictions. In specification (6), we look at the consequences of using OTHER PRESENT as an exclusion restriction, which however has no effect on the reporting process. Notice 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 underreport, but its interaction with the other variables in the reporting process also worsens the general behaviour of the model. Consequently, as it is the case with the other estimates, the effect of being an immigrant becomes more insignificant, although still negative.

6.3 Weighted *versus* unweighted regressions of property crime

The presented estimates so far are obtained utilising regression models that make use of appropriate weights that restore representativeness of our 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 models are very close to the parameters of the models that use weights, there is some support of correct specification of the model. The estimates are presented in specification (3) of Table 6 for the NB2-Logit and specification (7) of Table 5 for the MisProbit.

It is noteworthy that, apart from the coefficient of NORTH dummy, the weighted estimator produces estimates that are very close to the unweighted ones for both NB2-Logit and MisProbit. Moreover, it is evident that in general the coefficients of the unweighted regression are more precisely estimated. A remarkable difference however is that in the unweighted estimation the estimated immigration-crime differential is higher, in terms

Table 6 Robustness checks - NB2-Logit

	Other present 1 (1)		Other present 2 (2)		No weights (3)		10-25 (4)		Violent crime (5)	
	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$
IMMIGRANT	-0.475	0.189	-0.306	-0.094	-1.384***	0.739	-1.448	0.888	-1.053***	1.203
	(0.595)	(0.972)	(0.640)	(0.931)	(0.473)	(0.740)	(1.050)	(1.329)	(0.358)	(0.900)
AGE	-0.619***	0.910***	0.422***	-0.939***	-0.659***	0.976***	-0.388**	0.436**	-0.091***	-0.023
	(0.180)	(0.208)	(0.145)	(0.182)	(0.163)	(0.153)	(0.168)	(0.189)	(0.011)	(0.022)
AGE^2	0.009***	-0.014***	-0.007***	0.014***	0.009***	-0.015***				
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)				
MALE	1.462***	-1.325**	0.125	1.139	1.777***	-1.505***	0.270	0.96	-1.053	4.413
	(0.394)	(0.659)	(0.477)	(0.708)	(0.463)	(0.568)	(0.300)	(1.060)	(0.898)	(5.654)
WHITE	0.175	0.734	0.869	-0.525	-0.143	0.983	1.137	-0.532	0.073	0.933
	(0.662)	(1.191)	(0.611)	(0.974)	(0.656)	(0.943)	(0.990)	(1.521)	(0.594)	(0.770)
SOUTH	0.860	-1.024	-0.270**	0.992	1.130*	-1.501*	3.133	-3.937	0.27	-0.711
	(0.601)	(0.902)	(0.507)	(0.771)	(0.617)	(0.812)	(2.933)	(3.067)	(0.584)	(1.081)
MIDLANDS	0.161	0.519	0.694	-0.416	0.284	-0.035	-0.460	0.2	-0.121	-0.197
	(0.558)	(0.995)	(0.627)	(0.878)	(0.469)	(0.730)	(0.713)	(1.009)	(0.578)	(1.111)
NORTH	1.588**	-1.373	-0.201	1.903	0.970	-1.002	0.545	-1.469	0.977	-1.297
	(0.711)	(1.164)	(0.575)	(0.845)	(0.648)	(0.995)	(0.460)	(1.397)	(0.842)	(1.202)
TRUTHFUL						-1.144***		-2.868***		-4.992
						(0.353)		(0.722)		(20.032)
OTHER PRESENT		-0.677		-0.600*						
		(0.533)		(0.349)						
CONSTANT	7.572**	-11.470***	-5.849***	11.870***	8.068***	-11.612***	7.072	-6.238	2.188*	3.699
	(3.009)	(3.664)	(1.538)	(2.799)	(2.673)	(2.365)	(4.370)	(5.371)	(1.132)	(21.341)
Sample size	11,604		11,604		11,604		4,981		11,640	
Log likelihood	-2,314.29		-2,314.61		-4,073.06		-1,109.34		-2,232.35	
$\hat{\alpha}$	41.94***		41.99***		36.09***		21.93***		27.63***	
$\hat{p}$	0.448		0.464		0.422		0.189		0.561	

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

of magnitude, and statistically significant at 1% for both NB2-Logit and MisProbit. This might be the case because the unweighted estimator is more efficient, so that the immigration coefficient in the weighted estimation is less precisely estimated. Furthermore, as we have included the *ethnic-boost* data set, immigrants are over-represented in my sample. Thus, using weights that restore representative has as a result to attach lower weights to the immigration sample, which may induce differences in the estimated immigration status coefficient.

6.4 Criminal behaviour of young people (10–25 years old)

In Section 4 we stressed that, because of the demanding nature of the estimators used in this paper and the large concentration of our crime variable on the zero outcome, an attempt was made to keep sample size as large as possible. An alternative strategy would instead be to somehow increase the variation of our outcome variables by increasing the proportion of positives. Noticing that younger individuals are the ones who report the most crime, we increased the percentage of positives from 5.8% to 9.8% by keeping only respondents between 10 and 25 years of age. Note that this is not an arbitrary selection, as this is the age group of main interest of the CJS. Therefore, by restricting our sample to this age group, we briefly examine the differences in criminal behaviour between young natives and young immigrants.

The results are presented in specification (4) of Table 6 for the NB2-Logit model and specification (8) of Table 5 for the MisProbit model. Note that in this case we exclude Age^2 as it does not fit the data. We find that the estimates are relatively less precise, probably because of the smaller sample size, but the estimated immigration-crime differentials are still negative and fairly large in the NB2-Logit model, although insignificant once again.

6.5 Violent crime

Here we briefly investigate whether the link between immigration status and violent crime is also negative. The violent crime estimates are presented in specification (5) of Table 6 for the NB2-Logit and specification (9) of Table 5 for the MisProbit. Firstly, we find that for some unidentified reason, count data models of violent crime do not behave very well. The estimates of specification (5) correspond to the only maximum we managed to find. We had to exclude Age^2, as otherwise convergence seemed impossible. This model indicates that immigrants are considerably less involved in violent crime. However, these estimates are rather unreliable.

In the contrary, the MisProbit model behaves much better. We see that the estimates of violent crime are in general in line with the property crime ones[14]. In addition, we can notice that the predicted probability of reporting at least one of the committed violent crimes is fairly similar to the one of our baseline property crime model. Concerning the effect of the immigration dummy, it is again negative but less significant than for property crime. This result also holds for the models without exclusion, with OTHER PRESENT as exclusion restriction, and without using weights (these results are not presented here but are available on request). Hence, immigrants are slightly less crime prone than natives for both crime types.

7 Interaction terms

In this subsection, using interaction terms, we investigate whether the immigration-property crime relationship depends on the region of residence or on ethnicity. 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 instance, 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. Regarding ethnicity, 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 exhibit different criminal behaviour from natives of the same ethnic group, as the latter is better adapted in the British lifestyle. The results are presented in Table 7 and are briefly discussed. Note that the following results in general hold for the MisProbit model as well (available on request).

Regarding location, from specification (1) of Table 7 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 or in North commit significantly less property crime than natives in London or in North, respectively (p-values are 0.018 and 0.003 respectively). Actually, immigrants located in London are overall the least crime prone group, since they exhibit a significantly 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 immigrants located in South is the most crime-prone category. However, their involvement in crime, although higher, is not statistically different from the involvement of natives in South (p-value is 0.29). Note that immigrants in Midlands are not significantly less involved in property crimes than their native counterparts. Finally, we note that natives located in North commit significantly more crime than natives located in Midlands or London.

But what channels could possibly explain the results above? It might be, for instance, that immigrants integrate in London more easily than in other locations, because of better labour market opportunities and large 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, Butcher 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 et al. 2013). 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 (see, specifications (2) and (3)). First of all, from (2) we can see that although white immigrants and white natives are equally involved in property crime, non-white immigrants are significantly less involved in crime than non-white

Table 7 Interaction terms

| | By region | | By ethnic group | | | |
| | (1) | | (2) | | (3) | |
	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$	$\hat{\delta}$	$\hat{\eta}$
IMMIGRANT	-1.944**	1.539	-2.444***	1.663	-0.141	0.161
	(0.822)	(1.534)	(0.931)	(1.552)	(0.731)	(1.113)
AGE	-0.603***	0.976***	-0.692***	1.037***	-0.713***	1.053***
	(0.188)	(0.179)	(0.209)	(0.198)	(0.215)	(0.203)
AGE^2	0.008***	-0.014***	0.010***	-0.015***	0.010***	-0.15***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
MALE	1.452***	-1.408**	1.388***	-1.245*	1.367***	-1.219*
	(0.368)	(0.657)	(0.383)	(0.641)	(0.396)	(0.664)
WHITE	-0.384	1.451	-0.625	1.249		
	(0.817)	(1.432)	(0.773)	(1.248)		
BLACK					0.305	-1.162
					(0.773)	(1.257)
ASIAN&OTHER					0.021	-1.821
					(1.032)	(1.537)
MIXED					2.100	-1.710
					(1.316)	(1.889)
SOUTH	0.516	-1.042	0.730	-1.169	0.816	-1.562
	(0.629)	(0.928)	(0.687)	(0.977)	(0.717)	(1.024)
MIDLANDS	0.015	0.518	0.051	0.488	0.065	0.195
	(0.566)	(0.974)	(0.578)	(0.960)	(0.641)	(1.050)
NORTH	1.600**	-1.843*	1.563**	-1.518	1.636**	-1.868
	(0.631)	(1.057)	(0.642)	(0.973)	(0.697)	(1.135)
TRUTHFUL		-1.171**		-1.240***		-1.306***
		(0.461)		(0.463)		(0.455)
IMMIGRANT*SOUTH	3.022**	-3.116				
	(1.449)	(1.449)				
IMMIGRANT*MIDLANDS	1.259	-3.130				
	(1.387)	(1.387)				
IMMIGRANT*NORTH	-0.521	1.989				
	(1.102)	(0.631)				
IMMIGRANT*WHITE			2.351**	-1.582		
			(1.128)	(1.827)		
IMMIGRANT*BLACK					-2.467*	2.085
					(1.310)	(1.892)
IMMIGRANT*ASIAN&OTHER					-1.45	1.251
					(1.530)	(2.585)
IMMIGRANT*MIXED					-3.097	1.591
					(2.249)	(3.645)
CONSTANT	8.135***	-12.269***	9.886***	-13.521***	9.609***	-12.272***
	(3.216)	(2.992)	(3.582)	(3.450)	(3.357)	(3.276)
Sample size	11,604		11,604		11,604	
Log likelihood	-2,307.74		-2,311.02		2,308.49	
$\hat{\alpha}$	40.77***		41.18***		40.77***	
$\hat{p}$	0.456		0.433		0.420	

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

natives (at 1% level). Now from (3), comparing each ethnic group of immigrants with their native counterparts (and using appropriate Wald tests), it is noteworthy that Black immigrants are also less likely to commit a property crime than Black natives (significant at 5% level). 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 et al. 2010). Note also that the involvement of Black natives in criminal activities is not different from the involvement of all other groups. Thus, it seems that Black immigrants exhibit unobserved cultural characteristics associated with lower involvement in criminal activity than the other groups. Finally, note that there is no significant difference in crime involvement between the other two immigrant ethnic groups and their native counterparts.

8 Conclusions

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

Regression models for count and binary data that control for underreporting were developed and used, as underreporting is a major concern in crime self-reports. Given that some parametric conditions hold, these models allowed for consistent estimation of both the determinants of true criminal activity and the determinants of underreporting, using only data on observed reported crime.

The results of these models showed that there is substantial underreporting of criminal activity, but, if anything, immigrants tend to underreport by less than natives. Nevertheless, it was stressed that the coefficients of the reporting process of both NB2-Logit and MisProbit models must be treated with caution, since the reporting process can be also interpreted as in a Zero-Inflation (ZI) framework (that is, there are some individuals who never participate in criminal activities, while the rest of them may do so and therefore may commit crimes). However, we also developed the ZI-NB2-Logit model which disentangles ZI from underreporting. The estimates of the latter indicated that the probability of being an individual who never participates in the illegal sector is 0.64, while for those who do not belong to the ZI category, the probability to report a committed property crime is 0.39. The estimates of this model also indicated that immigrants are (insignificantly) less likely to belong to the ZI category, and that once controlling for ZI the effect of immigration status on reporting behaviour loses in both magnitude and significance but it still retains its sign.

Regarding the immigration-property crime link, the estimates of the crime process suggested that, controlling for reporting behaviour or/and ZI, 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 showed that it is actually quite robust. For example, the results of the unweighted models implied 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.

In the theoretical discussion we noted 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 we obtained here 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 face less favourable legal market opportunities than natives, perhaps because of human capital limitations, discrimination, difficulties of adjustment, cultural conflict, etc. (see, Algan et al. 2010). However, at the same time, immigrants may be more risk averse. As a result, they might be more responsive to potential punishment and other deterrent factors (Butcher 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 may be satisfied even with much lower economic outcomes relative to 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. Hence, 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 regions affect their criminal behaviour, or that different areas attract different types of immigrants. Finally, we interestingly found that, due to 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.

Endnotes

[1] For details on the survey design of the CJS refer to Hamlyn et 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.

[2] 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, Martinez and Lee 2000).

[3] In the contrary, it can be shown that imposing exclusion restrictions on the reporting process is not enough for identification of θ. 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 η. However, since in the current empirical study this information is not available, this possibility is not discussed further. For details refer to Papadopoulos (2011).

[4] Hausman et al. (1998) also allow for a probability of misclassifying a true zero as a one. In our application, this would be interpreted as over-reporting of crime. However, in order to be in line with our count data models, and since over-reporting is not very likely in our case, we have restricted this probability to zero. Papadopoulos (2011) presents results of the full misclassification model.

[5] Strictly speaking, if NB2-Logit is the true DGP, MisProbit cannot be, because if the process generating the data produces a NB2 distribution, then the binary information

cannot come from a Probit, but should be seen as a censored at one NB2. Estimation of such a model was attempted, however, convergence seemed impossible. Instead, a censored Poisson-Logit model was estimated giving results that are very similar to the MisProbit model. However, the Probit model is more standard theoretically and using the Poisson instead of the NB2 seems even less reasonable than using Probit instead of NB2. So, although if we accept that NB2-Logit is correctly specified, MisProbit cannot be, we assume that it only slightly deviates from the true DGP.

[6] Wooldridge (2007) defines this as selection on covariates. Formally, if s is the selection dummy taking value one if the individual participates and zero otherwise, $\mathbf{x}$ is the vector of explanatory variables affecting criminal behaviour, and ϵ is the error term in the crime equation, the selection is on covariates if $P(s_i = 1|\mathbf{x}, \epsilon) = P(s = 1|\mathbf{x})$.

[7] A limitation of this construction is that there may be some natives who replied that they have lived in the UK less than their whole life, just because they left the UK for a certain period of time. These people will be classified as immigrants, although they should be considered as natives, particularly if the period of staying outside the UK was very short. Nevertheless, this number is expected to be quite small, 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 estimated by the OECD (8.8% in 2003 and 9.3% in 2004).

[8] Actually, to achieve convergence we had to impose several exclusion restrictions from both processes, which leads to serious model misspecification, as most of the excluded variables arguably belong to both processes.

[9] Following Anderberg et al. (2013), we also attempted controlling for the effect of unemployment rate and the risk of apprehension at a local geographical level. Because the lowest geographical level available in our data is the Police Force Area (PFA), using data from the 2004 Annual Population Survey and 2004/05 Police Force Assessment published by the Home Office, we constructed the unemployment rate and the number of police officers per 1,000 capita, both at PFA level, and matched this information to our CJS data. Although number of police officers in an area is arguably important in deterring crime, it is unclear what this variable would capture because of reverse causality (high police force reduces crime but areas with high crime rates are assigned a higher police force power). Since there are 42 PFAs, this variables take on only 42 different values. However, we find that these variables have absolutely no explanatory power, neither in crime nor in the reporting process. Therefore, these variables are not used in the empirical analysis of this paper, but results are available on request.

[10] 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 TRUTHFUL on number of reported property crimes, showed that this is actually the case.

[11] If we instead keep WHITE as the base group and include the other three ethnic group dummies, our results (which are available on request) show that the immigration coefficient becomes -0.535. Moreover, BLACK and ASIAN&OTHER individuals are less crime prone than WHITE ones, but MIXED individuals are more crime prone. However, once again, all these differences are statistically insignificant.

[12] We tried different specifications, such as including a quadratic term of DEPRIVATION, or including this variable in dummies form. All results show that this variable has no explanatory power, not even without controlling for regional dummies or in the models that do not control for underreporting.

[13] For specifications (2) these numbers are 0.421 for natives and and 0.210 for immigrants, while for (5), 0.506 for natives and 0.240 for immigrants. Again, immigrants

commit almost half the crimes committed by natives. However, these differences are not statistically significant either.

[14]Note that the tetrachoric correlation coefficient is 0.576, so that it is not the case that the estimates are 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 only account for a very small proportion of the total number of property or violent crimes (1.2% for property crime and 1.1% for violence).

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

Appendix A: attitudes towards immigrants in the UK

In this appendix 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: "immigrants increase crime rates" (using a 5 points Likert-type scale, where 1 = "strongly agree" and 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.

These unconditional probabilities, 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 design (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 brief 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[15].

The results of an Ordinal Probit regression model are presented in the 3 columns of Table 8, where a simple dummy for year 2003 aims at capturing the evolution in respondents' attitudes. Note that the dependent variable is recoded such as value 1 now denotes "totally disagree ..." and value 5 denotes "totally agree ...". We include covariates for gender, age, education, political ideology (specification 1), region, union, marital and employment status (specification 2), and citizenship status (specification 3).
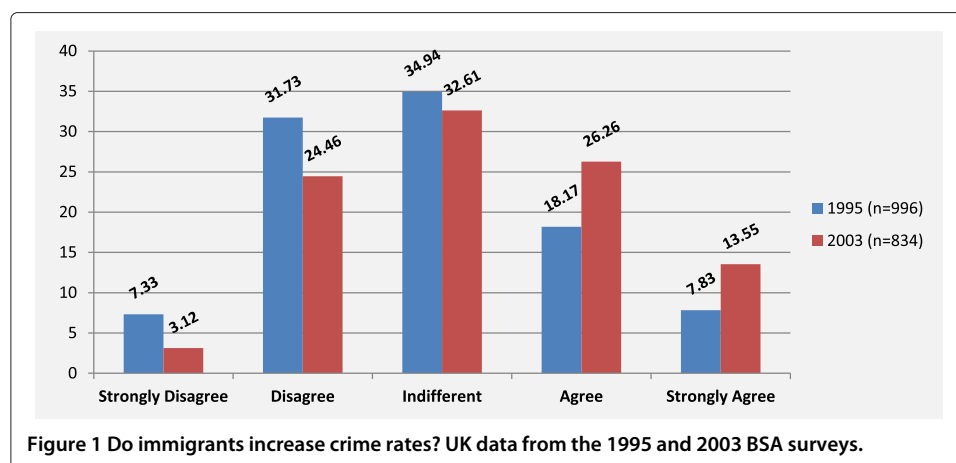

Figure 1 Do immigrants increase crime rates? UK data from the 1995 and 2003 BSA surveys.

Table 8 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 (*Both parents NOT citizens*)			
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.
Parentheses with *text* indicate the reference group.
*, **, *** denote 10%, 5%, 1% level of significance respectively.

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 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, while 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 does 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).

Appendix B: a model of participation in property crime

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, holding leisure constant, optimally decides how to allocate his available time, τ, between legal and illegal activities, denoted as τ_ℓ and τ_i respectively.

If the individual participates in the legal sector, he can be either employed (State A) or unemployed (State B) depending on the, exogenously given, probability of unemployment $\mu(m)$, where m is a binary indicator for immigration status. If employed, he receives wage $w(\tau_\ell, m)$ with $\frac{dw(\cdot)}{d\tau_\ell} > 0$, whereas if unemployed, he receives the unemployment benefit $D(\tau_\ell)$ with $\frac{dD(\cdot)}{d\tau_\ell} > 0$. It is also assumed that $\underline{w}(\tau_\ell) > D(\tau_\ell)$ and $\frac{dw(\cdot)}{d\tau_\ell} > \frac{dD(\cdot)}{d\tau_\ell}$, where $\underline{w}(\tau_\ell)$ is the minimum wage rate. On the other hand, if the individual participates in the illegal sector, he receives the criminal wage $k(\tau_i, m)$, which consists of financial and psychological outcomes measured in their monetary equivalent, with $\frac{dk(\cdot)}{d\tau_i} > 0$. Thus, psychological costs associated with crime, such as regret, uneasiness, etc, are incorporated in $k(\cdot)$. 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 crimes, he also faces the probability of apprehension, $\pi(\tau_i, m)$, with $\frac{d\pi(\cdot)}{d\tau_i} > 0$, and if apprehended, punishment, $P(\tau_i, m)$, occurs with certainty (without loss of generality), with $\frac{dP(\cdot)}{d\tau_i} > 0$. Punishment is also measured in its monetary equivalent and happens at the end of the period, so that the individual discounts it by a rate of $\rho(m)$.

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,

$$U(\tau_i, \tau_\ell) = (1 - \mu(m))\, u\,(y_a) + \mu(m)\, u\,(y_b) - \rho\ \pi(\tau_i, m)P(\tau_i, m), \qquad (12)$$

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 $u'(y_J) > 0$, and $u''(y_J) < 0$, where $J = (a, b)$. Henceforth, m is omitted from the equations for brevity. Thus, the individual needs to decide how to allocate his available time between legal and illegal activities in order to

Papadopoulos *IZA Journal of Migration* 2014, **3**:12
http://www.izajom.com/content/3/1/12

Page 28 of 30

maximize (12) subject to the time constraints, $\tau = \tau_i + \tau_\ell$, and, $\tau_i \geq 0, \tau_\ell \geq 0$. The Kuhn-Tucker first order conditions are,

$$\frac{dU(\tau_i)}{d\tau_i}\tau_i = 0, \quad \frac{dU(\tau_i)}{d\tau_i} \leq 0, \quad \tau_i \geq 0; \frac{dU(\tau_\ell)}{d\tau_\ell}\tau_i = 0, \quad \frac{dU(\tau_\ell)}{d\tau_\ell} \leq 0, \quad \tau_\ell \geq 0. \tag{13}$$

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

$$\left((1-\mu)u'(y_a) + \mu\, u'(y_b)\right)\frac{dk(\cdot)}{d\tau_i^*} - \left((1-\mu)u'(y_a)\frac{dw(\cdot)}{d\tau_\ell^*} + \mu\, u'(y_b)\frac{dD(\cdot)}{d\tau_\ell^*}\right)$$
$$= \rho\left(\frac{d\pi(\cdot)}{d\tau_i^*}P(\cdot) + \frac{dP(\cdot)}{d\tau_i^*}\pi(\cdot)\right), \tag{14}$$

so that the marginal utility obtained from criminal activities minus the marginal utility obtained from legal activities must be equal to the marginal punishment. Since the RHS of 14 is weakly positive, the individual will spend time on illegal activities *iff* the marginal utility from criminal activities is at least as high as the marginal utility from the legal sector. This is the marginal compensation required to cover for the risk of spending time on committing crimes.

As the criminal wage rate is in general small compared to the legal wage rate for most property crimes, and if we consider that for most people the criminal wage further decreases by psychological costs, the corner solution where someone allocates all his time in legal actions is highly likely. 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 very high psychological costs for most people. On the other hand, the individual will specialise in the illegal sector, *iff* the marginal utility from the legal activities plus the marginal cost of punishment is smaller than the marginal utility from illegitimate activities, which is highly unlikely.

What could 14 and a simple comparative statics analysis tell us about differences in the criminal activities between a typical immigrant and a typical native? We notice that immigration status affects criminal behaviour through many channels, as *m* appears in most determinants of 14. Firstly, starting from an equilibrium where the individual participates in crime, an increase in the marginal utility gained from the legal sector will decrease the LHS of 14 and therefore, *ceteris paribus*, participation in crime becomes less likely, 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 statics analysis shows that $\frac{d\tau_i^*}{d\mu} > 0$ *iff*, $\frac{u'(y_b)}{u'(y_a)} > \frac{(dk/d\tau_i - dw/d\tau_\ell)}{(dk/d\tau_i - dD/d\tau_\ell)}$. 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 strict concavity of $u(\cdot)$. Thus, the LHS of this inequality is higher than one and this inequality always holds. Now, since immigrants on average face lower legal opportunities, such as lower $\frac{dw(\cdot)}{d\tau_\ell}$, or higher μ, we would expect immigrants to be more crime prone than natives. Regarding criminal opportunities, there is no evidence on whether immigrants exhibit a higher or a lower $\frac{dk(\cdot)}{d\tau_i}$ than natives.

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 14. As discussed in Section 2, we would expect that

the average immigrant faces higher $\pi(\tau_i)$ and $P(\tau_i)$, and therefore 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 their utility functions, are quite important on determining criminal behaviour. For example, people that are very "patient" discount potential punishment less heavily (higher ρ) which increases the RHS of 14. Moreover, more risk averse individuals are represented by "curvier" 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 14 becomes smaller). In both cases, a higher marginal compensation is required to cover for the extra risk. Thus, because discount factors and risk attitudes may be quite different between immigrants and natives, we expect their participation in criminal activities to be different as well. 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.

References
Alesina A, La Ferrara E (2005) Ethnic diversity and economic performance. J Econ Lit 43:762–800
Algan Y, Dustmann C, Glitz A, Manning A (2010) The economic situation of first and second-generation immigrants in France, Germany and the United Kingdom. Econ J 120:F4–F30
Anderberg D, Rainer H, Wadsworth J, Wilson T (2013) Unemployment and domestic violence: theory and evidence. IZA Discussion Paper Series, No. 7515. http://ftp.iza.org/dp7515.pdf
Aquilino WS (1993) Effects of spouse presence during the interview on survey responses concerning marriage. Public Opin Quart 57:358–376
Beine M, Docquier F, Rapoport H (2008) Brain drain and human capital formation in developing countries. Econ J 118:631–652
Bell B, Machin S (2013) Immigrant enclaves and crime. J Regional Sci 53:118–141
Bell B, Fasani F, Machin S (2013) Crime and immigration: evidence from large immigration waves. Rev Econ Stat 95:1278–1290
Bianchi M, Buonanno P, Pinotti P (2012) Do immigrants cause crime? J Eur Econ Assoc 10:1318–1347
Borjas GJ (1999) Immigration and welfare magnets. J Labor Econ 17:607–637
Borjas GJ (2003) The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market. Q J Econ 118:1335–1374
Budd T, Sharp C, Mayhew P (2005) Offending in England and Wales: first results from the 2003 crime and justice survey. Home Office Research Study 275
Butcher FK, Piehl AM (1998) Cross-city evidence on the relationship between immigration and crime. J Policy Anal Manag 17:457–493
Butcher FK, Piehl AM (2007) Why are immigrant' incarceration rates so low? Evidence on selective immigration, deterrence, and deportation. NBER Working Paper, No. 13229. http://www.nber.org/papers/w13229.pdf
Cameron AC, Trivedi PK (1998) Regression analysis of count data. Cambridge University Press, New York
Card D (2009) Immigration and inequality. Am Econ Rev 99:1–21
Dustman C, Fabbri F, Preston I (2005) The impact of immigration on the British labour market. Econ J 115:F324–F341
Ehrlich I (1973) Participation in illegitimate activities: a theoretical and empirical investigation. J Polit Econ 81:521–565
Feilzer M, Hood R (2004) Differences or discrimination: minority ethnic young people in the youth justice system. Youth Justice Board, London
Feller W (1968) An introduction to probability theory and its applications, Volume 1, Edition 3. John Wiley, New York

Gourieroux C, Monfort A, Trognon A (1984) Pseudo maximum likelihood methods: theory. Econometrica 52:681–700

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

Hart HH (1896) Immigration and crime. Am J Sociol 2:369–377

Hausman JA, Abrevaya J, Scott-Morton FM (1998) Misclassification of the dependent variable in a discrete-response setting. J Econometrics 87:239–269

Jaitman L, Machin S (2013) Crime and immigration: new evidence from England and Wales. IZA J Migr 2:19

Junger-Tas J (1997) Ethnic minorities and criminal justice in the Netherlands. In: Torny M (ed) Ethnicity, Crime, and Immigration Crime Justice, vol. 21. University of Chicago Press, pp 257–307. http://www.press.uchicago.edu/index.html

Junger-Tas J, Marshall IH (1999) The self-report methodology in crime research. Crime Justice 25: 291–367

Lochner L, Moretti E (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. Econ J 112:85–106

Manacorda M, Manning A, Wadsworth J (2012) The impact of immigration on the structure of wages: theory and evidence from Britain. J Eur Econ Assoc 10:120–151

Martinez R, Lee MT (2000) On immigration and crime. Criminal Justice 1:486–524. US Department of Justice, Office of Justice Programs

Ousey GG, Hubrin CE (2009) Exploring the connection between immigration and violent crime rates in U.S. Cities, 1980–2000. Soc Probl 56:447–473

Papadopoulos G (2011) Immigration and crime: a microeconometric study. Unpublished PhD Manuscript, Department of Economics, University of Essex. https://ueaeprints.uea.ac.uk/41424/1/thesis_GP.pdf

Papadopoulos G, Santos Silva JMC (2008) Identification issues in models for underreported counts. Department of Economics, University of Essex, Discussion Paper Series, No. 657. http://www.essex.ac.uk/economics/discussion-papers/papers-text/dp657.pdf

Sharp C, Budd T (2005) Minority ethnic groups and crime: findings from the offending, crime and justice survey 2003. 2nd Edition, Home Office Online Report 33/05. http://www.clydebankhigh.org.uk/New%20CHS%20Website/Files/modern%20studies/Adv%20Higher/CausesEffects%20of%20Crime/Articles-handouts/Minorities%20and%20offending.pdf

Smith DJ (1997) Ethnic origins, crime, and criminal justice in England and Wales. In: Torny M (ed) Ethnicity, crime, and immigration. Crime Justice, vol. 21. University of Chicago Press, pp 101–182. http://www.press.uchicago.edu/index.html

Spenkuch JL (2013) Understanding the impact of immigration on crime. American Law and Economics Review. doi: 10.1093/aler/aht017

Staub KE, Winkelmann R (2013) Consistent estimation of zero? Inflated count models. Health Econ 22:673–686

Tonry M (1997) Ethnicity, crime, and immigration. In: Torny M (ed) Ethnicity, crime, and immigration. Crime Justice, vol. 21. University of Chicago Press, pp 1–29. http://www.press.uchicago.edu/index.html

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL (1998) Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. Science 280:867–873

Wadsworth 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. Soc Sci Quart 91:531–553

Winkelmann R (2008) Econometric analysis of count data. 5th ed. Springer-Verlag, Berlin

Winkelmann R, Zimmermann KF (1993) Poisson logistic regression. Department of Economics, University of Munich, Working Paper No. 93–18

Wooldridge JM (2007) Inverse probability weighted estimation for general missing data problems. J Econometrics 141:1281–1301

Wooldridge JM (2010) Econometric analysis of cross section and panel data. 2nd ed. MIT Press, Cambridge

Yeager MG (1997) Immigrants and criminality: a review. Crim Justice Abstr 29:143–171