ECONOMIC ANALYSES OF CRIME IN ENGLAND AND WALES

By

LU HAN

A thesis submitted to the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Economics
Birmingham Business School
College of Social Science
The University of Birmingham
September 2009

ABSTRACT

This thesis includes three empirical studies detecting the determinants of crime in England and Wales. We firstly apply time series analyses to look for cointegrating relationships between property crimes and unemployment as well as law enforcement instruments. We extend our study by employing panel data and corresponding techniques to control for area-specific fixed effects as well as the endogeneity of law enforcement variables. In our third study, we allow crime rate to have spatial spillover effect, in other words, the crime rate in one area is affected by, in addition to its local crime-influential factors, the crime rates and crime-related factors in its neighbouring areas. We demonstrate this result by constructing a theoretical model and testing it by applying spatial analysis regressions. Our main findings can be summarized as follows: First, property crimes are better explained by economic models of crime than violent crimes. Second, law enforcement instruments always have negative effects on both property and violent crimes, indicating their deterrence and incapacitation effects as predicted. Third, social-economic factors, such as unemployment and income level, have two effects on property crimes: opportunity and motivation. Their net effects on property crime rates depend on the type of crime as well as the time period being examined. And finally, there is indeed spillover effect existing in crime rate. For burglary, theft and handling, and robbery, the crime rate in one area is positively and significantly correlated with the crime rates from its neighbouring areas. Furthermore, the crime rate of sexual offences of one area is negatively related to such crime rates in neighbouring areas.

To Grandma, Grandpa, Mum and Dad—My Eternal Love

ACKNOWLEDGEMENTS

First and foremost, my heartfelt thanks to my supervisors, Dr. Siddhartha Bandyopadhyay and Prof. John Fender, whose sound guidance, unyielding support made this research process an intellectually stimulating experience for me. I am deeply grateful for their professional and dedicated supervision.

I also owe my genuine gratitude to Prof. David Maddison, Mr. Ralph Bailey and Dr. Samrat Bhattacharya for providing me with invaluable comments and advise on the technical aspects of econometrics. Their support is crucial for the completion of the thesis. Likewise, I am indebted to many other staff members in the Department of Economics at the University of Birmingham. They create such an inspirational and motivating environment in which I feel honoured to have studied over the years. My thanks also go to the supporting staff at the department, who always promptly deal the 'small matters' with a 'big smile', from room booking to toner replacing.

My appreciation goes to friends and colleagues, who are always here with laugh and laughter to make this long research process much more interesting.

Needless to say, this thesis would not have been possible without unparalleled support and immense encouragement from my family. My beloved parents not only have made this work possible by financially sponsoring me with their hard-earned money, they also have constantly been there to offer me great strength to endure the 'roller-coaster rides' throughout the demanding research process. Therefore, above all, I give my greatest gratitude to mum and dad, to whom this thesis is dedicated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

ADF	Augmented Dicky-Fuller Test
BP	Breusch and Pagan Test
CD	Compact Disc
ECM	Error Correction Model
GHET	Groupwise Heteroskedasticity
GMM	Generalised Method of Moments
GIS	Geographic Information System
GWR	Geographically Weighted Regression
ILO	International Labour Organisation
KB	Koenker and Bassett Test
LR	Log Likelihood Ratio Test
ML	Maximum Likelihood
OLS	Ordinary Least Squares
PP	Phllips-Perron Test
SAR	Spatial Autoregressive Model
SatNav	Satellite Navigation System
SMA	Spatial Moving Average Model
TSLS	Two-Stage Least Squares
U.K.	The United Kingdom
U.S.	The United States

Chapter One: Introduction

1.1 BACKGROUND AND AIMS

The aim of this thesis is to understand the determinants of crime rates in England and Wales by applying various empirical analyses. Becker (1968) and Ehrlich (1973) provided a framework for the economic analysis of crime by assuming potential offenders are economically rational and utility maximizing. Becker and Ehrlich predicted that crime should be deterred by tougher law enforcement policies because higher "price" for committing crimes should discourage the incentives to participate in illegal activities of rational offenders. Further, worse labour market opportunities for potential offenders will encourage their criminal activities by reducing their opportunity cost of doing so. Following the theories of Becker and Ehrlich, many scholars were interested in empirically testing the predicted relationships between crime and variables which influence the potential criminal's rational decision by affecting the costs and benefits of committing crime. However, there are various econometric issues (for example endogeneity) in most of these papers and often significant crime influencing variables are not included in the estimation leading to an omitted variable bias.

I use a broad range of crime types, both property and violent, and adopt multiple techniques to analyse this issue in England and Wales. In this thesis, I aim to provide a relatively comprehensive understanding on the intriguing but complex relations between different crimes and various social factors through multiple types of empirical analysis.

Such analysis has distinct implications for policy analysis. The effective response to crime is one of the essential priorities faced by any government. However, in order to provide effective crime control, it is crucial to clearly understand the interaction between crime and various factors. Therefore, the empirical analyses in this thesis could supplement to the existing literature of crime in England and Wales to analyze the determinants of criminal activities. Furthermore, as implied in Becker (1968), combating crimes through law enforcement and justice system is costly to social welfare. If, by empirical analysis, it could be demonstrated that crime is also heavily affected by certain social-economic factors, there could be an additional method to control crime which is less costly and provide other long term benefits to society. For example, higher legal income level is predicted to reduce crime through increasing offenders' opportunity cost of committing crimes; similarly, a decrease in unemployment rate could also reduce crime by improving the labour market opportunities for potential criminals (Ehrlich 1973). If various empirical tests in England and Wales can constantly confirm the validation of such predictions, it might be able to draw the implication that economic growth could be an effective way to ease the issue of crime. Comparing to law enforcement effort, combating crime through economic growth could be less costly because economic prosperity itself is the target of both government and population. Reduced crime rate becomes a positive "side-effect" of better economic conditions. Moreover, different types of crimes differ in their nature and thus their reactions to the crime influential factors. Empirical testing could reveal such diversity and, as a result, a better understanding of the nature of each individual type of crime may be achieved. We may realize that different policies may be better suited for combating different kinds of crime.

1.2 THESIS STRUCTURE

The thesis is divided as follows:

In Chapter Two, we provide a background by surveying the economic literature on crime starting with Becker (1968) who first formalized how crime could be looked upon as any other economic activity. A lot of theoretical and empirical work has tried to extend Becker’s theory and have used empirical analogues to Becker’s model in trying to determine the effects that different factors have on crime rates. I survey some of the important works in that areas; most of which are based on the cases in the U.S.

In Chapter Three, we start our empirical analysis by applying time series data and techniques to look at both long-run and short-run relationships between crime rates and a limited number of variables. We use national level time series data over the period 1971-2000 (for the overall crime and fraud and forgery, the examined period is 1971-1997) to study the correlations between property crimes and their related factors of law enforcement instruments and unemployment rate. The aim of this chapter is to identify the unemployment – crime relationship, which is famous for being ambiguous in sign, by examining their temporal variations. In addition, we also test the predicted negative correlations between property crimes and law enforcement variables from the angle of time series analysis. The modern econometric techniques enable us to test the stationarity of different variables as well as long-term cointegration and short-term error correction models. Our results mainly suggest that, in long-run, the overall and property crime rates are cointegrated with unemployment and law enforcement instruments. While unemployment has positive cointegration with overall crime, burglary, and theft and handling, it has negative cointegration with fraud and forgery. The custody rate, as one of the law enforcement variables, is negatively cointegrated with both overall and property crimes. Detection rate, the other law enforcement variable, has negative cointegration with overall crime and positive cointegration with each individual property crimes. In short-run, on the other hand, the change in custody rate constantly and negatively affects each crime rate being examined, the change in unemployment is positively correlated with the overall crime, burglary, as well as theft and handling.

We extend our analyses in Chapter Four by adding violent crimes into our analyses. More importantly, we introduce a relatively complete set of crime-related explanatory variables in our empirical model, such as law enforcement, social-economic conditions, demographic composition, as well as once-lagged crime rate reflecting a dynamic pattern in crime rate. Our panel data disaggregated by 43 police force areas in England and Wales over the period 1992-2005 enables us to study the relationships between different crime rates and their influential variables based on the information that varies by both areas and years. In addition, such a data structure also provides the capacity to control for area-specific fixed effects that would be otherwise correlated with the independent variables. We adopt generalized method of moments (GMM) technique, in addition to OLS and fixed effects models, to eliminate the area-specific fixed effects and to apply instrumental variables for endogenous law enforcement variables as well as once-lagged crime rate. Our main findings are: firstly, property crimes are better explained by our empirical model than violent crimes. Individual crime types could have different response to the explanatory variables. Secondly, the law enforcement variables broadly show negative correlations with different crime rates confirming their deterrence (and probably also incapacitation) effects on crime as predicted by theories. Thirdly, the social-economic factors, such as unemployment and real earnings, mainly pick up their opportunity (i.e. the opportunities for such crime) effects on property crimes indicating worse social-economic conditions will reduce property crimes due to fewer opportunities available. And finally, each crime rate being analysed shows significant and positive correlation with its once-lagged value. This result suggests strong persistence in crimes and we will give detailed explanations in this chapter.

In Chapter Five, we examine whether there are spatial spillover effects in crime rates. We relax the assumption adopted by most works in this area (including our previous analyses) that crime rate in one area is only affected by its local related factors. Instead, we allow the crime rate of each area to depend on, in addition to local factors, the crime rates and explanatory variables in neighbouring areas. The idea stems from the fact that, if criminals are mobile, policy in one area could affect the crime rate in neighbouring areas as well. For example, toughened law enforcement in one area could drive criminals to spillover into neighbouring areas causing a negative externality. Alternatively, relatively affluent crime opportunities in one area could attract the potential criminals spillover from neighbouring areas and leaving their "home" areas with lower crime rates. Based on the assumption of spill-over effect, we firstly construct a simple theoretical model containing two regions contiguous to each other. By showing that the number of crimes spilling over from one area to the other is affected by the relative cost and opportunities between the two regions, we derive several predictions for later tests. Accordingly, we constructed two empirical models: in the first one, we allow the crime rate in one area to be affected by a set of local explanatory variables as well as the crime rates from its neighbouring areas; in the second model, we assume the crime rate in one area is predicted by not only its local related factors, but also those factors from neighbouring areas. Our empirical models are tested by applying spatial lag and spatial error models, as well as a spatial contiguous matrix, on panel data disaggregated by police force areas over the period 1998-2001. Our application of panel data in spatial analysis models is a major improvement comparing to other spatial analysis papers because, in this way, we are able to explicitly include both area-specific and year-specific fixed effects that would be otherwise correlated with the independent variables. Our results suggest that there is indeed spatial spillover effect between the crime rates of neighbouring areas specifically for burglary, theft and handling, robbery and sexual offences. In addition, among the explanatory variables we included, the strongest predictors for the crime rate in one area are its local detection rate and real earnings. While detection rate mostly has negative effect on crime, real earnings are positively correlated with property crimes reflecting its opportunity effect.

Finally, we conclude our main findings in Chapter Six where we also discuss the limitations in this work as well as provide prospects of future research in this field.

Chapter Two: Literature Review

2.1 Introduction

In this chapter, we survey some important literature in the field of economic analysis of crime. Becker (1968) and Ehrlich (1973) start analysing the phenomenon of crime in economic frameworks and most of later works construct their models accordingly. By assuming potential criminals are economically rational and utility maximizing, Becker and Ehrlich theoretically relate crime rate to various factors such as law enforcement effort and social-economic status. Furthermore, they predict that an increase in either the probability of apprehension or severity of punishment will reduce people's incentives to commit crimes; either higher illegal payoffs or lower legal returns will increase one's participation in criminal activities. Moreover, unemployment rate, measuring the risk of legal labour market, has ambiguous effect on crime. Many later papers are particularly interested in empirically testing these predictions made by Becker and Ehrlich.

We construct this literature review by putting papers with the same aim together and hope to provide a general background for the literature analysing crime with economic theories. The papers most related to our empirical studies will be introduced and discussed in each individual chapter. In this chapter, section two introduces classic economic frameworks applied on crime, while section three discusses the papers aiming to detect the deterrence effect of law enforcement. We talk about literature identifying the well-known unemployment – crime relationship in section four and examine the papers testing the effects of the overall labour market conditions in section five. And finally, section six focuses on detecting the effect of the proportion of young people.

2.2 ECONOMIC THEORIES OF CRIME

Becker (1968) can be seen as the first work which formally analyses the phenomenon of crime in an economic model and its special contribution is undeniable. Almost every paper later on has cited Becker (1968) in a significant position as the original inspiration. This paper is no exception.

Becker (1968) looks at the issue of crime control from the angle of social welfare. As increasing the probability of apprehension and the severity of punishment are both costly, it targets to identify the optimal levels of punishment by minimising the social loss induced by both crime combat and crimes themselves.

As a part of the model, the damage from offences *(D)* has been related to the number of offences *(O)* and can be written as the following equation:

$$D(O) = H(O) - G(O),$$

where H denotes the harm to the victims and the society and G denotes the gain to offenders. Therefore, the net damage to the society is simply the difference between the two. It is reasonable to assume that both H and G are increasing with the number of offences O as long as each additional crime will cause positive harm to the victim or/and the society and positive gain for the criminal himself. In mathematical term, that is $H'(O) > 0$ and $G'(O) > 0$. Therefore, the sign of $D'(O)$ depends on the relative magnitudes of H' and G'.

The costs of apprehension and punishment should both be a positive function of the level of offences *O*. The cost of apprehension can be written as:

$$C = C(pO),$$

where *p* represents the probability of apprehension and *O* represents the level of offences. Both *p* and *O* should be positively related to the cost of apprehension *C*: $C_p > 0$ and $C_O > 0$. The social cost of punishment depends on the exact form of the punishment. This problem can be simplified by imposing a coefficient *b*:

$$f' \equiv bf$$

where *b* takes the value of 0 for fines and greater then unity for torture, probation, parole, imprisonment and most other punishment. In this way, the punishment *f* imposed on the criminal can be transferred to the social cost f'. For punishment taking the form of fines, the coefficient *b* can be quite close to 0 if the money transfer can be regarded as costless. On the other hand, punishments of other forms are costly not only to the criminals but also to the society. Therefore, the induced social cost should be greater than the cost on the offender and the coefficient *b* is greater than 1.

Then, Becker derives a function relating the number of offences one would commit his probability of apprehension, his severity of punishment and other relevant variables such as his income level in both legal and illegal activities, his frequency of nuisance arrests and his willingness to commit offences etc. This function can be expressed as:

$$O_i = O_i(p_i, f_i, u_i),$$

where O_i is the number of offences person *i* would commit; p_i and f_i represent his probability of apprehension and severity of punishment respectively; u_i represents all other relevant variables. An increase in either p_i or f_i would reduce one's expected utility from an offence and thus would tend to reduce the number of offences he would commit. This point can be easily proved by the expected utility function of a potential offender

$$EU_i = p_i U_i(Y_i - f_i) + (1 - p_i) U_i(Y_i),$$

where Y_i represents his income from an offence; U_i represents his utility function; f_i represents his monetary equivalent of the punishment. The change in the expected utility with respect to the probability of apprehension and the severity of punishment can then be derived by taking first order conditions:

$$\frac{\partial EU_i}{\partial p_i} = U_i(Y_i - f_i) - U_i(Y_i) < 0$$

and

$$\frac{\partial EU_i}{\partial f_i} = -p_i U_i^{'}(Y_i - f_i) < 0 .$$

Both first order conditions would be negative as long as the marginal utility of income is positive. This is saying that the expected utility from an offence would be decreasing as the probability of apprehension and severity of punishment increase. Thus, the number of offences one would commit, O_i, should have the following properties:

$$O_{p_i} = \frac{\partial O_i}{\partial p_i} < 0$$

and

$$O_{f_i} = \frac{\partial O_i}{\partial f_i} < 0 .$$

Literally, the number of offences one would commit would be negatively related to both the probability of apprehension and the severity of punishment.

As the concern is normally the aggregated level of offences, it can be derived by summing all the O_i. Its determinants, though, need some minor corrections. The aggregated number of offences, *O*, would still be affected by the probability of apprehension, the severity of punishment and other relevant factors. However, such determinants are likely to differ from one person to another. To tackle this issue, Becker takes the average values for *p*, *f*, and *u*, for simplicity. The aggregated number of offences can then be expressed as:

$$O = O(p, f, u),$$

where *p*, *f* and *u* are denoting the average values of p_i, f_i and u_i. This function is expected to have the same properties as the individual function: the aggregated number of offences would be negatively correlated with both *p* and *f*. Furthermore, it would be more responsive to the change in *p* than to the same change in *f* if, and only if, most offenders are risk-lover.

The aggregated supply of offences function in Becker (1968) contains the probability of apprehension, the severity of punishment and a third factor *u*. Although, Becker mentions that *u* represents a set of relevant variables such as one's incomes in legal and illegal activities, family background, education, preference for risk etc, he does not explicitly analyse the effect of any of these relevant factors. This gap is filled by Ehrlich (1972; 1973). In both works, one is assumed to be able to spend his time on either legal or illegal activities or both. His time allocation depends on the relative expected utility from each activity. While Ehrlich (1972) only provides a verbally analytical model, Ehrlich (1973) has formally developed a mathematical framework analysing the participation in illegal activities with a choice under uncertainty theory.

An essential assumption in Ehrlich (1973) is that one is free to combine a number of legitimate and illegitimate activities or switch occasionally from one to another during any period throughout their lifetime. His object is to maximize his expected utility by optimally allocate his time and other resources between legal and illegal activities. For simplicity, the optimal participation in illegal activities is analysed in a one-period uncertainty model. It is assumed that there is no training or other entry costs required in either legal and illegal activities and neither are there costs of movement between the two. Since activity *l* (legal activity) is safe, its net returns are given with certainty by the function $W_l(t_l)$, where *t* denotes the time input. On the other hand, activity *i* (illegal activity) is risky in the sense that its net returns are conditional upon two states of the world: *a*, apprehension and punishment with probability p_i, and *b*, getting escaped with probability of $1 - p_i$. If successful, the offender obtains a payoff $W_i(t_i)$ which is his gain from the illegal activity net his costs of inputs. If apprehended and punished, his payoff will be reduced by an amount of $F_i(t_i)$. Therefore, one would obtain either

$$X_b = W' + W_i(t_i) + W_l(t_l)$$

With probability $1 - p_i$, or

$$X_a = W' + W_i(t_i) - F_i(t_i) + W_l(t_l)$$

With probability p_i, where W' denotes the market value of the individual's assets. Accordingly, his expected utility is given by

$$EU(X_s, t_c) = (1 - p_i)U(X_b, t_c) + p_i U(X_a, t_c)\,.$$

The problem has now become maximizing the above equation with respect to the choice variables t_i, t_l and t_c subject to a time constraint

$$t_0 = t_i + t_l + t_c .$$

The solutions require that

$$t_i \geq 0 , t_l \geq 0 \text{ and } t_c \geq 0$$

It is not difficult to show that, given the amount of time allocated to consumption t_c, the optimal allocation of working time between activity *i* and *l* must satisfy the first order condition

$$-\frac{w_i - w_l}{w_i - f_i - w_l} = \frac{pU^{'}(X_a)}{(1-p)U^{'}(X_b)},$$

where $w_i = \frac{dW_i}{dt_i}$, $f_i = \frac{dF_i}{dt_i}$ and $w_l = \frac{dW_l}{dt_l}$. Apparently, a necessary prerequisite is that the potential marginal penalty, f_i, should exceed the differential marginal return from illegal activity, $w_i - w_l$, because otherwise the marginal opportunities in *i* would always dominate those in *l*.

As the two-state-of-the-world assumption seems less realistic, the model can be expanded by incorporating more variables which implies more states of the world. For example, unemployment rate, as an important indicator of the legal labour market, can be included to measure the risk in legal activities while p_i still measures the risk in illegal activities. The preceding analysis can still be applied in this four-state-of-the-world condition. In order to obtain an interior solution, the following first-order condition should be satisfied:

$$\begin{aligned}&(1-p_i)(1-u_l)U_a^{'}(w_i - w_l) + (1-p_i)u_l U_b^{'}(w_i) + \\ &p_i(1-u_l)U_c^{'}(w_i - f_i - w_l) + p_i u_l U_d^{'}(w_i - f_i) = 0\end{aligned},$$

where u_l represents the probability of unemployment in legal activities *l*; *a*, *b*, *c* and *d* are the four relevant states of the world.

Both equations of first-order conditions have identified the basic factors determining entry into and optimal participation in illegal activities. Furthermore, they have some implications on the effects of relevant factors on the participation in illegal activities. Firstly, an increase in either p_i or f_i, with other variables being constant, would reduce one's incentive to enter and participate in illegal activities because it increases the expected marginal cost of punishment $p_i f_i$. Secondly, an increase in the marginal or average differential return from illegal activity $w_i - w_l$, resulting from either an increase in illegal payoffs or a decrease in legal incomes or both, would increase one's incentive to enter into or allocate more time to illegal activities. Thirdly, an increase in the probability of unemployment, u_l, has a more ambiguous effect on the incentive to participate in illegal activities. This is because an increase in the probability of the least desirable state of the world (unemployed in legal activities and failed in illegal activities) would increase one's demand for wealth in this state and might decrease his incentive to participate in illegal activities. However, the partial effect of an increase in u_l would unambiguously be positive on his incentive to participate in illegal activities. And finally, a decrease in the amount of time allocated to nonmarket activities t_c would not affect one's relative preference between different states of the world and therefore lead to a positive scale effect on participation in activities *i* and *l*: more time would be spent in both legal and illegal activities.

Given the validity of the preceding analyses, an individual supply-of-offenses function can be specified

$$q_{ij} = \varphi_{ij}(p_{ij}, f_{ij}, w_{ij}, w_{lj}, u_{lj}, \pi_j),$$

where *i* and *l* still denote illegal and legal activities; *j* denotes individual *j*; q_{ij} denotes the number of offenses committed by individual *j*; and π_j denotes a set of other variables that might affect the number of offenses committed by individual *j*.

If all individuals are identical, the aggregated supply-of-offenses functions can be easily derived given the individual function. As mentioned in Becker (1968), the relevant variables are likely to differ from person to person due to personal heterogeneity. Ehrlich takes the average values of these variables to incorporate in the aggregated supply of offenses function

$$Q_i = \Psi_i(P_i, F_i, W_i, W_l, U_l, \Pi_i),$$

where Q_i represents the aggregated supply of offenses; P_i, F_i etc. represent the average values of p_{ij}, f_{ij} etc.; and Π_i includes, in addition to those environmental variables, all the moments of the distributions of *p*, *f* etc. other than their means.

Based on early theoretical frameworks, Ehrlich (1996) has described a "market model" which is also based on the assumption that offenders, like other members of a society, respond to incentives. Such incentives, both negative and positive, do not largely differ from those introduced in Ehrlich (1973): negative incentives are referring to the factors deterring potential offenders to commit crimes such as detection and punishment; positive incentives are the factors that encourage legal activities as an alternative to committing offenses, such as the employment and earning opportunities, rehabilitation programs and lower disparity in the income distribution.

The most important difference between this model and previous ones is probably that the equilibrium is achieved through the interactions between three parties instead of two: in addition to offenders and the law enforcement authorities, the potential victims also affect the equilibrium. The potential offenders are still behaving to maximize their expected returns from both legal and illegal activities given the incentives they are facing. Potential victims are expected to desire protection from offenses, such as insurance against crime, burglar alarm systems, safety deposit boxes etc. However, such protections come at a cost. The optimal level of protection, therefore, depends on a position where the marginal cost equals the marginal benefit of the protection. In this way, the level of private protection, measured by the expenditure, can be related to crime rate which is the proxy for the probability of becoming a victim. In addition, the expenditure on private protection can also affect the cost of offenses for potential offenders. For law enforcement authorities, their targets will still be maximizing the social welfare by minimizing the loss from offenses. The optimal level of public expenditure on law enforcement can be found by balancing the marginal cost and the marginal benefit by imposing the law enforcement instruments. The expenditure on law enforcement would not only affect the potential offenders by changing their expected punishment, but also affect the demand for private protection against offenses. The supply of offenses, together with the demand for private and public protections against offenses, forms the basic components of the “market model”. The equilibrium is reached when the level of offenses is such that neither offenders, potential victims, nor the government feel the need to adjust their behaviours.

This "market model" has derived some implications concerning the behaviours of different parties. Firstly, crime is a persistent and "normal" social issue which has survived through history regardless of the prevailing economic, political or social system. This is because both private and public protections against offenses are costly and certain level of offenses would lead to socially optimal equilibrium. Secondly, some social, political or demographic conditions affect the level of offenses. For example, an economic growth and real asset accumulation could raise the potential payoffs to offenders in many illegal activities. Thirdly, changes in the court decisions or sentencing guidelines over time would shift the "tax" paid by offenders if convicted and thus their net returns.

2.3 DETECTING THE EFFECT OF LAW ENFORCEMENT

The preceding part has introduced the theoretical models analysing the relationship between the level of offenses and its influential factors. The validity of these theoretical frameworks, however, rests on the assumed deterrent effect of law enforcement. The idea is that law enforcement—apprehension and punishment of offenders—serves partly to deter future offenses by increasing the expected cost of breaking the law for both actual and potential offenders. Given the increased expected cost of committing offenses, both actual and potential offenders would have lower incentives to participate in illegal activities. Such assumptions, however, have been seriously questioned in the criminological literature of the past hundred years or so. As stated in Ehrlich (1972), "an extreme view is that law enforcement has no deterrent effect on offenders (at least those who commit serious crimes), essentially because offenders are very different from other human beings." In fact, criminological literatures usually claim that the criminal behaviours should mainly be attributed to the offenders' unique motivations which, in turn, have been linked to his unique "inner structure" or the impact of exceptional social and family circumstances. Therefore, demonstrating the existence of the deterrent effect of law enforcement is not only important for the theoretical frameworks, but also meaningful for the later literature which are developed upon these theories. Ehrlich (1972) is such a work that presents empirical evidence to support the deterrent effect of law enforcement.

In the theoretical model developed in Ehrlich (1973), the factors affecting one's participation in illegal activities include his legal and illegal returns, the probability and severity of punishment, and the probability of being unemployed in legal labour market. A negative relationship between the law enforcement and the level of offenses is predicted accordingly. However, the probability and severity of punishment may affect crime rate through a different channel other than deterrence when the punishment takes the form of imprisonment. This is called the preventive effect, also known as the incapacitation effect. The preventive effect takes place when the convicted offenders are punished by sentencing into prison: as they are separated from the potential victims, they can be prevented, at least temporarily, from committing further crimes.

Although both deterrent and preventive effects of law enforcement are expected to reduce the level of offenses, it is still essential to establish an independent deterrent effect of law enforcement. This is because, firstly, preventive effect only exists for the offenders who are punished imprisonment. Furthermore, the preventive effect only works temporarily up to the release of the prisoner. The deterrent effect, on the other hand, should influence both actual and potential offenders and does not depend on the form of punishment. Secondly, reducing the level of offences through preventive effect induces extra cost. Sentencing offenders into prison requires expenditure on relevant facilities and supervision personnel. In addition, imprisonment restricts the prisoners to contribute to national production.

Theoretically, the deterrent effect of law enforcement is distinguishable from the preventive effect when the punishment takes the form of imprisonment. It can be shown that the preventive effects of one percent increase in the probability of apprehension P or the number of periods offenders actually serve in prison T are approximately the same after a period of adjustment to the dynamic processes generated by the changes in P and T. The preventive effect, caused by the change in either P or T, is equal to the ratio

$$\sigma = \frac{PT}{1+PT}.$$

Since σ is positively related to PT, the expected length of imprisonment, the preventive effect of law enforcement might be relatively small for less serious offenses. The derived preventive effect of law enforcement implies a feasible method to distinguish between the preventive and deterrent effects: as the preventive effects of P and T are expected to be equal in magnitudes, significant difference between the estimated elasticity of crime rate to the changes in these variables would suggest the existence of an independent deterrent effect of law enforcement. Furthermore, the ratio σ is supposed to be lower than unity even if P is close to unity. The existence of deterrent effect might also be detected if the empirical coefficients of P and T are close to, or even greater than, unity.

Ehrlich (1972) carries out empirical analysis to test his theory introduced above. The analysis is based on cross-state analysis of seven “index crimes” punishable by imprisonment. The years being covered are 1940, 1950 and 1960. The empirical model takes the number of offenses per capita as dependent variable and the independent variables include the ratio of the number of commitments into prison to the number of offenses, the average time served in prison, the median family income, the percentage of families below one half of the median family income, and the percentage of non-while population. The estimated results have exhibit a remarkable consistency with the theoretical predictions. Specifically, while relevant variables are controlled, the coefficients of the probability and length of imprisonment are consistent with the hypothesis that law-enforcement instruments have deterrent effects on offenders. Such deterrent effect is independent from the preventive effect of imprisonment.

Levitt (1996) focuses on the relationship between the prison population size, measuring the severity of punishment, and different crime rates. This paper is motivated by the phenomenon that the incarceration rate in the United States has more than tripled over two decades: from 1970s to 1990s. Meanwhile, the rate of imprisonment in the United States is three to four times greater than most European countries. Such high level of incarceration, however, does not seem to be accompanied by obvious declines in crime rate. Therefore, this papers aims to identify the effect of prison population on crime rates when other relevant variables are being controlled.

Given the trends of both prison population and crime rates, one cannot simply conclude that increased incarceration rate has been ineffective as a law enforcement instruments. Crime rates are also affected by other determinants such as the labour market opportunities, potential offenders' family and educational backgrounds, gang involvement etc. In order to generate relatively precise estimate for prison population, other relevant variables should be controlled as completely as possible. Furthermore, simultaneous bias could also affect the estimation. Increased incarceration is likely to reduce crime due to both deterrent and incapacitation effects. However, increased crime rate could lead to larger prison population through either higher probability of apprehension or more severe punishment or both. Consequently, the OLS estimation will be biased by such reverse causality. Thus, Levitt chooses to employ instrument variables which should be correlated with the endogenous variable, prison population, and uncorrelated with crime rates. The instrument variable employed in this paper is the status of state prison overcrowding litigation. Unsurprisingly, the existence of overcrowding litigation reduces the growth rates of prison population.

The data used in this paper is state-level panel data over the years 1971-1993. The model under estimation is given by the equation

$$\Delta\ln(crime_{st}) = \beta\Delta\ln(prison_{st-1}) + X_{st}^{'}\theta + \gamma_t + \varepsilon_{st},$$

where the subscript s and t represent the states and years respectively; $crime_{st}$, as the dependent variable, represents the per capita crime; $prison_{st-1}$, as the main concern of this paper, represents the once-lagged per capita prison population; X_{st} is a set of covariates including the per capita income, unemployment rate, the percentage of population who are black, the percentage population who live in metropolitan areas and a set of age distribution variables; year-specific effect is captured by the parameter γ_t.

When treating prison population as exogenous, its coefficient is -0.099 for violent crime and -0.071 for property crime. After applying instrument variables for prison population, its effect on violent crime is almost four times greater than before. Similarly, the effect of prison population on property crime, when being intrumented, is four times higher than before. Such results strongly support the effectiveness of imprisonment, as law enforcement instrument, on crime. Furthermore, the same estimation is applied on seven individual crime categories and has generated consistent results: in all seven cases, applying instruments lead to more negative coefficient for prison population. In addition, assault, robbery and burglary are more responsive to the change in imprisonment.

According to the classic models developed in both Becker (1968) and Ehrlich (1973), an increase in either the probability of apprehension or the severity of punishment would tend to reduce the expected returns from illegal activities and hence the crime rate. Levitt (1996) has confirmed the negative effect of prison population size, measuring the severity of punishment, on crime rate. In Levitt (1998), effort has been devoted to identifying the effect of increased arrest rate on crime while the arrest rate is assumed to measure the probability of apprehension. Furthermore, this paper also designs a strategy trying to distinguish the deterrence and incapacitation effects of increased arrest rate.

The approach distinguishing between the deterrence and incapacitation effects lies on specific assumptions: firstly, criminals commit multiple offenses and do not specialize in one particular type of crime; secondly, certain types of crime are substitutes for each other while other types of crime are not. As long as these two assumptions are satisfied, the deterrence and incapacitation effects of arrest rate could be separately identified. This is because an increase in the arrest rate for any crime would lead to a reduction in all crimes due to the incapacitation effect. In contrast, if criminals are rational and different crimes are substitutes for one another, the deterrent effect implies that increasing the arrest rate for one crime would lead to a decrease in own crime, but an increase in other crimes as criminals would substitute away from the own crime. Therefore, the expected sign of deterrence effect depends on the relationship between different types of crimes: deterrence is negative for the own-crime rate and positive for the substituting crime rates. For the non-substituting crimes, the deterrence is not expected to have significant effect. If the incapacitation effect exists, on the other hand, an increase in the arrest rate for one crime would not only reduce own-crime rate, but also reduce the substituting-crime rates since more criminals have been locked behind bars.

In order to support the assumption that criminals commit multiple crimes which belong to different types, strong evidence has been found: surveys of prisoners in the United States reveal that the median number of non-drug crimes committed in the year preceding their most recent arrest is twelve to fifteen. Furthermore, there is evidence suggesting that the majority of criminals are in fact generalists. Only one in twenty released murderers who recidivate will have his next arrest to be for murder.

The corresponding number for released robbers and auto thieves is less than one in three. For burglars and larcenists, the number is slightly less than one in two.

For the purposes of empirical analysis, it is assumed that violent crimes are substitutes for each other, and property crimes are substitutes for each other. However, there is assumed to be no substituting relationship between violent and property crimes. The empirical analysis has been carried out based upon the annual-level panel data which covers 59 of the largest U.S. cities over the period 1970 to 1992. The estimated equation is given by

$$\begin{aligned}\ln c_{it} &= \alpha_i + \lambda \ln(a_{it}^o / c_{it}^o) + \theta \ln(a_{it-1}^o / c_{it-1}^o) + \varphi_s \ln(a_{it}^s / c_{it}^s) + \\ &\kappa_s \ln(a_{it-1}^s / c_{it-1}^s) + \varphi_n \ln(a_{it}^n / c_{it}^n) + \kappa_n \ln(a_{it-1}^n / c_{it-1}^n) + X_{it}^{'}\beta + \varepsilon_{it}\end{aligned},$$

where the subscripts *i* and *t* denote cities and time periods respectively; the subscript *o* denotes own-crimes, *s* and *n* refer to the crimes that are substitutes and non-substitutes for crime *o*; variables *c* and *a* represent the number of reported crimes and the arrests respectively and the ratio between them represents the arrest rate. The estimated equation is describing a model in which the crime rate is assumed to be determined not only by the contemporary and once-lagged arrest rates of its own crime type, but also by the contemporary and once-lagged arrest rates of both substituting and non-substituting types of crimes. Furthermore, a set of covariates is also expected to influence the crime rate. Such covariates include city population size, the percentage of black population, the percentage of population residing in female-headed households, the percentage of population between the age 15 and 24, the unemployment rate, the combined local spending on education and public welfare etc. The estimated results can be summarized by the following four points. Firstly, a negative relationship between crime rate and own arrest rate has been found for six out of seven types of crime, murder is the only exception. Secondly, the arrest rates of presumably non-substituting crimes should only affect the own crime rate through the incapacitation effect. More specifically, the coefficients of the non-substituting arrest rates are expected to be negative. The estimation has shown that, for all seven categories, the estimated coefficients are indeed negative, although the coefficients are significant only in the cases of robbery and rape. Thirdly, the arrest rates of substituting crimes are expected to generate ambiguous coefficients: the incapacitation effect would negatively influent the crime rate under study while the deterrence effect would have positive influence. The net effect, therefore, would be ambiguous. In practice, the estimated coefficients of substituting arrest rates are mixed: negative and significant in three cases, positive and significant in two cases and insignificant in two cases. Given the obtained results, one important implication for crime control policies is that both deterrence and incapacitation effects are indeed existing and distinguishable. The optimal level of crime control, therefore, should be decided while incorporating this factor.

As demonstrated by the papers introduced previously, the probability of apprehension and the severity of punishment have both been confirmed for their deterrent effects on crime rate. One practical problem that usually rises during the empirical analysis is that both factors, the probability of apprehension and the severity of punishment, can have different proxies as their counterparts in the empirical analysis. For example, the probability of apprehension is traditionally measured by either detection rate or conviction rate. The counterpart of severity of punishment, on the other hand, has more choices. Given the relevant literatures, the most commonly used variables measuring the severity of punishment are the prison population size and the average length of imprisonment. One would be reasonable to suspect that death penalty, as an extreme form of punishment, would have some deterrent effect on crime. As more convicted criminals get sentenced to death penalty, the potential criminals would rationally expect tougher punishment once convicted and hence become more cautious when making the decision on whether or not commit crimes. However, the paper Levitt (2003) has cast some doubts on the deterrent effect that death penalty is expected to have. As stated in this paper, “a number of studies have found evidence supporting a deterrent effect of the death penalty (Cloninger, 1977; Deadman and Pyle, 1989; Ehrlich, 1977; Ehrlich and Liu, 1999; Layson, 1985; Mocan and Gittings, 2001). A far larger set of studies have failed to find deterrent effects of capital punishment (e.g., Avio, 1979, 1988; Bailey, 1982; Cheatwood, 1993; Forst, Filatov, and Klein, 1978; Grogger, 1990; Leamer, 1983; Passell and Taylor, 1977).” Aside from these mentioned literatures, Levitt has found some statistical records which may support the claim that the deterrent effect of death penalty is far too limited than expectation. In 1997, totally 74 prisoners were executed in the United States and it has been the highest amount in thirty years. However, the executed prisoners only take approximately 2 percent of the total inmates under death sentence up till the end of 1997. Even among those who have eventually put to death, there is a long lag between sentencing and execution. Therefore, as claimed in this paper, given the high discount rates of many criminals and the fact that many homicides are committed by individuals under the influence of alcohol or drugs, it is hard to believe that punishment with such a long delay would be effective.

Instead, the paper argues that the quality of life in prison is likely to have a greater impact on criminal behaviour than the death penalty. More specifically, the lower the quality of life in prison, the greater the punishment for a fixed amount of time served. Furthermore, poor prison conditions, unlike death penalty, would affect all inmates regardless of the crimes committed. Levitt (2003), therefore, aims to test the deterrent effect of prison conditions on crime using panel data from the United States.

The data is a panel of annual, state-level observations covering the continental United States for the time period 1950-1990. The dependent variables being analysed are three types of crimes: murder, violent crime (excluding rape) and property crime (excluding larceny)[1]. The explanatory variables of primary interest are the execution rate and the death rate among the prisoners from all sources other than execution. The death rate is incorporated in the analysis as a proxy of the prison conditions. In addition to the prisoner death and execution rates, a range of criminal justice, economic and demographic variables are also included. More specifically, the certainty of severity of a state's criminal justice system is proxied by the number of prisoners per violent crime and the ratio of prisoners to state population. Given that the contemporary values of these two variables would probably cause endogenous biases to the estimation, their once-lagged values have entered the estimated equation instead. The economic statuses are measured by the real state per capita income and the insured unemployment rate. The demographic controls include the percentage of black population, the percentage of metropolitan population, age distribution variables and infant mortality rate.

The main findings of this paper can be summarized as follows. Firstly, in all estimations with homicide as the dependent variable, prison death rate is negative and

[1] Rape is excluded because the data was not collected until 1957. Larceny is omitted because the important changes in its definition over the time period examined.

precisely estimated. The decline in homicides associated with one additional prison death varies from -0.1 to -0.8 across specifications. In contrast, the coefficient of execution rate is extremely sensitive to the choice of specification and has much larger estimated standard errors. The proxies of criminal justice system have generated different results: the prisoner-per-crime variable significantly loses its impact on homicide as the specification includes the full set of independent variables; the prisoner-per-capita variable, on the other hand, becomes increasingly negative as more independent variables come on board. Higher income is consistently associated with higher homicide rate while higher unemployment rate has the opposite effect. The effect of larger fraction of black population is, as expected, positive on murder rate and the effect of more metropolitan residents is surprisingly negative. The age distribution and infant mortality variables do not generate significant coefficients constantly.

Secondly, the death rate has obtained negative and significant coefficients in all estimations with violent crime as the dependent variable. According to the results, the elasticity of violent crime with respect to prison death rate varies from -0.05 to -0.17. On the other hand, no systematic effect of execution has been found for violent crime. The prisoner-per-crime variable, in this case, has shown much more significant effect comparing to the prisoner-per-capita variable, and this result is opposite to that obtained for homicide. While the economic factors appear to be weakly associated with violent crime, both fractions of black and urban residents exhibit positive association with violent crime in all specifications. Furthermore, the population under the age 25 is positively correlated with violent crime rate only in specifications with a limited set of control variables.

Thirdly, the prison death rate has lost some significance in explaining property crime: its coefficient is significant in only a few specifications. The execution rate, as usual, has generated ambiguous coefficient which is quite sensitive to the specification. Both the number of prisoners per crime and the number of prisoners per capita are negatively correlated with property crime. Higher unemployment rate, as expected, is associated with higher property crime and such positive correlation has also been found for the fraction of black people and the population under the age of 25.

In addition to the basic empirical model upon which the previously discussed results are based, expanded models have also been estimated in order to test the robustness of the results. The first extension to the basic model is constructed by allowing for prison death and execution rates to have both contemporaneous and lagged effect on crime rates. The second method is splitting the data set into two parts using the year 1971 as the break point. The last attempt is using logarithm, instead of level, of the observations to estimate the basic empirical model. Generally speaking, the extended models do not generate significant evidence that is against the deterrent effect of prison death rate found by the basic model. Furthermore, the execution rate has again obtained coefficient sensitive to the choice of specification.

As most empirical literature have devoted their effort in investigating the determinants of crime in the United States, other countries have not been completely left out although the amount of their relevant literature is remarkably smaller. Wolpin (1978) is one of papers studying the criminal behaviour in England and Wales. As stated in the paper, “crime of almost every variety has increased enormously over the past 80 years in England, at the same time the risk of capture and severity of punishment have declined.” Hence, this work attempts to understand the interrelationship between crime and its determinants during the time period 1894-1967.

Wolpin derives a supply-of-offenses function based on the theoretical frameworks introduced in Becker (1968) and Ehrlich (1973). In addition to empirical analysis, Wolpin has also extended the classic theoretical models by incorporating more factors measuring the law enforcement. Traditionally, the law enforcement effort is measured by the probability of apprehension and the severity of punishment. The probability of apprehension is proxied by either the detection rate or conviction rate. The severity of punishment, on the other hand, is usually measured by either prison population or average length of imprisonment. This work, however, has incorporated the probabilities of detection, conviction and punishment at the same time. Furthermore, different forms of punishment have also been included in the expected utility function and they are expected to have separate effects on crime. As individuals are assumed to maximize their expected utilities, interesting implications can be made regarding the elasticity of expected utility with respect to the law enforcement variables. Firstly, the elasticity of expected utility with respect to the probability of apprehension, conviction and punishment can be ordered. More specifically, the one percentage increase in the probability of apprehension would lead to the greatest decrease in the expected utility, comparing to the same percentage changes in the probability of conviction and punishment. Meanwhile, the change in punishment would have the smallest effect on the expected utility. Secondly, when the punishment only takes three forms: imprisonment, recognizance and fines, with imprisonment being the most severe punishment, one percent increase in the imprisonment rate would reduce the expected utility most, the recognizance rate next, and the fine rate last.

The empirical analysis is carried out for the time period 1894 to 1967 excluding the war years, 1914-1919 and 1939-1945. The supply-of-offenses equation is firstly estimated under the assumption that the law enforcement variables are not affected by the level of crime. In other words, the law enforcement variables are taken as exogenous at the first stage. The dependent variables under estimation are 8 types of crime rates.[2] The explanatory variables can be divided into two groups: group one contains a set of law enforcement variables such as the clearance rate, conviction rate, imprisonment rate, recognizance rate, fine rate and average length of sentence; groups two includes variables that mainly reflect social-economic and demographic status.[3]

In most cases, the estimated coefficients of law enforcement variables are negative. The major exception is the average length of sentence: its estimated coefficient is inconsistent with respect to both "sign" and "significant". This result could possibly be due to the weak relationship between the average length of sentence and the actual time served. The other law enforcement variables, generally speaking, perform better in the estimations for property crimes. The predicted order of elasticities is partially verified with only the conviction rate being estimated with quite weak elasticity. For violent crimes, the law enforcement instruments perform less well: the estimated

[2] The dependent variables include larceny, burglary, robbery, auto theft, malicious wounding, felonious wounding, all offenses against the person and all offenses.

[3] These controlling variables are the percentage of males aged from 10 to 25, the unemployment rate, the real weekly wage in manufacturing for manual workers, the real per capita GDP, the per capita number of police, real per capita local government expenditure, the proportion of those arrested for indictable offenses given legal aid and a dummy variable with value 0 before WW II and value 1 thereafter.

elasticities are less precise and the coefficient of imprisonment rate is far too large to conform to the predicted elasticity-ordering.

With respect to alternative forms of punishment, same percentage increases in the probability of imprisonment, recognizance and fine are, in most cases, consistent with the prediction: a change in the probability of imprisonment would yield the biggest response of crime rate while the same change in the probability of fine would lead to the smallest response.

The impacts of other controlling variables can be summarized by several points. Firstly, young males have been confirmed to have greater propensity to engage in property crimes although they have not been proved to be more likely to commit violent crimes. This finding could be attributed to the fact that juveniles would be treated with much milder punishment once convicted. Therefore, if they respond to the incentives and opportunities the same way as adults, their lower expected cost of punishment would increase their offense rate. Secondly, the unemployment rate is positively correlated with the overall and property crime rates. This is consistent with the implication derived in Ehrlich (1973) that, although the net effect of unemployment is ambiguous, its partial positive effect on crime rate is definite because higher unemployment rate would reduce the opportunity cost of committing crimes from the legal labour market. Thirdly, the proportion of individuals aged 15 and over attending school is negatively correlated with crime rates of both violence and property. The reason could be that increasing educational attainment reduces the necessity to solve disputes with violent actions and hence the violent crime rate. More directly, school attendance reduces the time available for other activities, including committing crimes. However, the negative correlation between school attendance and crime rates could be running from the opposite direction: criminal records could reduce future employment opportunities and, therefore, reduce the expected return from the schooling investment. Fourthly, the degree of urbanization positively affects the aggregate crime rate. For individual type, increasing urbanization raises the levels of larceny and burglary but reduces malicious wounding and auto theft. Fifthly, the per capita GDP, supposed to measure the average gain from crime, has negative effect for property crimes and positive effect for violent crimes. Furthermore, the real weekly wage for manual workers, a measure of the alternative legal wage, has obtained the same result: negatively correlated with property crimes and positively correlated with violent crimes. According to the theoretical models, increasing the average gain from crime is predicted to increase the property crimes as the expected return from such crimes has gone up. On the other hand, increasing the alternative legal wage is expected to reduce property crime by increasing the opportunity cost from legal labour market. The estimated results, therefore, imply that the per capita GDP has failed to capture the motivating effect of higher expected gain from property crimes.

The previously discussed results are based on the assumption that all the law enforcement variables are exogenous to the estimation system. Such assumption, however, could be easily challenged. On the one hand, crime rate could be reduced by increased law enforcement investment through deterrence and incapacitation effects; on the other hand, increased crime rate could reduce the productivities of law enforcement instruments by sharing the resources. Therefore, the seemingly negative relationship between crime rate and law enforcement variables runs in two opposite directions and such simultaneous interaction would cause estimation biases with ordinary estimation methods. This paper has adopted a three-equation system to deal with the simultaneous issue and generate unbiased estimation. In addition to the supply-of-offenses equation, it is assumed that the conviction rate and the per capita number of police are endogenous to the system, both of which are functions of crime rate and other relevant variables. [4] The estimations of the determinants of conviction rate and police demand have yielded encouraging results. However, as the author stated, "more detailed data sets and the resolution of several conceptual problems are needed in order to gain further insights into crime-prevention decisions and the validity of the social-loss framework."

2.4 DETECTING THE EFFECT OF UNEMPLOYMENT

Opportunities from legal labour market can be regarded as alternative options to both potential and actual offenders. As implied in Ehrlich (1973), an individual is assumed to allocate his available time between legal and illegal activities in order to maximize his expected utility. The optimal time allocation is thus depending on the expected opportunities from both legal and illegal markets. As alternative to criminal activities, increased legal opportunities, such as higher probability of employment and expected income, are supposed to increase the expected return from legal labour market and, in such case, individuals are expected to allocate more time to engage in legal activities. As predicted in Ehrlich (1973), an increase in the expected legal income would reduce the number of offenses one would commit by increasing the opportunity cost of

[4] Conviction rate is assumed to be a function of crime rate, per capita number of police, the proportion of defendants given legal aid, imprisonment rate, lagged conviction rate, per capita number of registered motor vehicles as well as all the environmental variables used in the supple-of-offenses equation. The per capita number of police is defined as a function of its lagged value, lagged crime rate, the per capita number of registered motor vehicles and local expenditures.

spending time on illegal activities. However, the effect of unemployment rate, measuring the uncertainty in legal labour market, would be more ambiguous to predict. Ehrlich (1973) has pointed out a partial positive effect that unemployment rate could have on the number of offenses given that higher unemployment rate would reduce the expected opportunities from legal labour market. Meanwhile, however, higher unemployment would also increase the probability that one would end up with the lease-desired situation: unemployed in legal market and failed in illegal market. In order to avoid such situation, an individual is likely to allocate less time on illegal activities.

In addition, Cantor and Land (1985) also develop a theoretical model which has predicted two distinct and counterbalancing effects of unemployment rate on crime. The analysis is based on the argument, proposed in Cohen and Felson (1979b), that the production of conventional crimes requires the presence of *a.* motivated offenders and *b.* suitable targets in *c*. the absence of effective guardians. Therefore, an increase in crime should be caused by an increased convergence of both motivated offenders and suitable targets under the situation of ineffective guardian. Following this proposition, the motivation effect of unemployment as predicted in Ehrlich (1973) only reflects the positive relationship between unemployment rate and the number of potential offenders. On the other hand, Cantor and Land (1985) has also predicted an opportunity effect of unemployment on crime which reflects a negative correlation between unemployment rate and potential victims.

As argued in Cantor and Land (1985), higher unemployment rate would reduce the number of suitable targets for property crimes due to two reasons. Firstly, by removing from employment, more people would be staying in or around their residing neighbourhoods. Such people and their properties are at reduced risks of becoming victims. Secondly, higher unemployment rate could be taken as a signal of declining economy. Therefore, producing and consuming activities are also likely to slow down as a result for both employed and unemployed. Such reduced property accumulation would thus provide fewer opportunities for property crimes. Given the reasons stated above, all other things being equal, higher unemployment rate would lower the probability of concurrence between motivated offenders and careless targets, and hence, lower the property crime rates.

Same logic can be applied on violent crimes because, contrary to the image promoted by the media, a substantial fraction of violent crimes involve causal acquaintances or strangers. According to the victimization survey in the United States, 60 percent of rapes and aggravated assaults reported to the police involve total strangers. In addition, 76 percent of murders were committed by offenders other than families and friends: 26.4 percent by non-family member; 13.3 percent by stranger; and 35.8 percent by people unknown to the police. Consequently, higher unemployment rate may reduce violent crimes through the same channel as property crimes: by reducing the availability of potential victims.

This paper, Cantor and Land (1985), has also designed an empirical model to test the proposed motivation and opportunity effects of unemployment rate on crime and applied such model on the time series data set of the United States covering the time period 1946-1982. The equations under estimation are specified as

$$\Delta C_t = \alpha + \beta_1 U_t + \beta_2 \Delta U_t + \varepsilon_t$$

$$\Delta^2 C_t = \alpha + \beta_1 U_t + \beta_2 \Delta U_t + \varepsilon_t$$

$$\Delta \ln C_t = \alpha + \beta_1 \ln U_t + \beta_2 \Delta \ln U_t + \varepsilon_t$$

$$\Delta^2 \ln C_t = \alpha + \beta_1 \ln U_t + \beta_2 \Delta \ln U_t + \varepsilon_t .$$

The dependent variables are either the differenced[5] levels or differenced logarithms of the index crimes: non-negligible homicide, forcible rape, aggravated assault, robbery, motor-vehicle theft, burglary, and larceny theft. The independent variables include both contemporary and differenced unemployment rate (in the form of either level or logarithm). The contemporary unemployment rate, both logged and unlogged, is supposed to capture the opportunity mechanism: once become unemployed, individuals would concentrate their activities in their residences and residential neighbourhoods and hence reduce their probability of victimization through the guardianship effect. Furthermore, as a signal of economic downturn, contemporary unemployment rate is assumed to immediately reflect people's declining consuming behaviour and thus reduce the availability of suitable targets. The differenced unemployment rate, on the other hand, is incorporated to represent the motivation impact which is less likely to be contemporaneous. It is argued in this paper that newly unemployed people are usually covered by unemployment benefit from government and therefore would not be immediately under the pressure of financial crisis. After a while, as benefits and other financial supports decline or even stop, unemployed individuals might be more likely to engage in criminal activities as a solution to financial crisis.

[5] The first differencing is to eliminate a linear secular trend; the second differencing is to eliminate a quadratic trend.

The estimated results, in general, have provided mixed support for the theoretical expectations: five out of seven index crimes have been detected to have small but significant correlation with unemployment rate. Furthermore, while the negative contemporaneous effect is indicated for all crimes expect rape and assault, the only evidence of a lagged motivational effect is for crimes that have a property component.

The methodology adopted in Cantor and Land (1985) is, with no doubt, an innovation in separating the opportunity and motivation effects that unemployment might have on crime rate. However, this strategy has been criticised by several following papers. For example, Greenberg (2001) has raised various questions concerning different aspects of Cantor and Land (1985) and its extension Land *et al.* (1995). First of all, using the differenced unemployment rate to capture the motivation effect is inappropriate. It is argued, in Greenberg (2001), that it is unrealistic to assume that large fraction of unemployed individuals would have enough savings or benefits to cover their finance for a while after becoming unemployed. In fact, many would face serious financial difficulties very quickly after losing a job. Therefore, annual lag of unemployment rate would be insufficiently fine-grained to detect the change in financial status over a much shorter time period. Furthermore, even if the motivation effect is indeed lagged, it is should be the once-lagged unemployment rate to enter the equation, rather than the differenced unemployment rate.

Secondly, Greenberg (2001) has questioned the approach adopted in Cantor and Land (1985) that only the dependent variables have been taken first or second order difference in order to remove time trend. It is a standard procedure that time trend should be eliminated before further analysis by taking differences. However, as claimed in Greenberg (2001), it is mathematically unacceptable if the differencing procedure is only carried out on the crime rates but not the explanatory variables.

The paper Hale and Sabbagh (1991) has also found the approach adopted in Cantor and Land (1985) is questionable mainly from the aspect of empirical approach. As broadly accepted, the first step of time series analysis is testing the stationarity of incorporated variables. Simply speaking, if the relevant variables are of different order of integration (i.e. they need different times of differencing to become stationary), ordinary estimation techniques would break down and generate unreliable inferences. Such problem exists in the strategy suggested in Cantor and Land (1985). If both crime rate and unemployment rate are integrated of order 1, they are non-stationary in levels but becoming stationary once first-differenced. Taking the first estimated equation for example

$$\Delta C_t = \alpha + \beta_1 U_t + \beta_2 \Delta U_t + \varepsilon_t ,$$

the model is actually trying to explain a stationary variable (differenced crime rate) with a non-stationary variable (unemployment rate) and another stationary variable (differenced unemployment rate). This approach is statistically invalid and the model is therefore mis-specified. As the same problem exists in every estimated equations proposed in Cantor and Land (1985), any conclusion drawn from this paper is probably unreliable concerning the relationship between unemployment and crime rates.

As Cantor and Land (1985) is trying to separately identify the opportunity and motivation effects that unemployment could have on crime, there are other papers only interested in estimating the net effect of unemployment rate on crime. For example, Fleisher (1963) employs time series data set trying to indentify the effect of unemployment on juvenile crimes. It is argued in the paper that higher unemployment rate not only creates difficulties for new entrants in the labour market in the sense of satisfying the desire for market goods and maintaining an acceptable living standard, it also make it harder for families to provide market goods and services for their children. Thus it is expected that unemployment could be positively correlated with crimes among young people.

In order to test the prediction discussed above, the paper has employed time series data structure which covers the period 1932-1961 (excluding the war time 1942-1945). According to the author, the time series data structure has its own merit in analysing the relationship between unemployment and crime. First of all, time series data reflects more clearly the trends of different variables over time. By covering relatively longer time period, it offers the opportunity to study the long-run relationship between relevant variables. Secondly, time series analysis could avoid incorporating regional differences in income, population characteristics and taste etc. which might produce disturbances that are hard to account for.

The dependent variable being estimated is the arrest rate for property crimes and this variable is expressed as the number o f arrests divided by the age-specific population. The independent variables include the male unemployment rate for ages 14-19 and 20-24, the total number of personnel in the United States army services, the ratio of property crime arrest rate for all ages to the rate of property offenses known to the police and a dummy variable splitting the whole time period into two parts with the year 1951. With OLS estimation, the results do support the prediction that the effect of unemployment on juvenile crime is positive and significant.

Some papers prefer to employ panel data analysis because such data structure could enrich both sample size and information which the analysis is based on. In addition, panel data could reflect the variations both over time and cross regions. The paper Raphael and Winter-Ebmer (2001) is one of the examples that investigate the relationship between unemployment and crime using panel data. More specifically, the data is disaggregated on state-level in the United States covering the years 1971-1997. The model under estimation is specified by the equation

$$crime_{it} = \alpha_t + \delta_i + \varphi_i time_t + \omega_i time_t^2 + \gamma unemployment_{it} + \beta X_{it} + \eta_{it} ,$$

where i and t index states and years. In addition to the unemployment rate, the model also incorporates a set of controlling variables which are represented by the matrix X. such controlling variables include alcohol consumption per capita, average income per worker, proportions of state residents that are black, living in poverty and residing in metropolitan areas, as well as prison population per 100,000 state residents. Furthermore, the model also includes state-specific and year-specific effects to eliminate the influence of factors that vary by either state or year. The state-specific linear and quadratic time trends are included to eliminate the variation in within-state crime rates caused by factors that are state-specific over time.

The specified model is firstly estimated with OLS regression as the basic analysis. As argued in the paper, however, the causal relationship could run from crime to unemployment. This is because higher crime rates could discourage employment growth and drive away existing firms and thus contribute to a state's unemployment rate. Alternatively, former criminals could find it difficult to participate in legal labour market given the criminal record and thus have to remain unemployed. In order to deal with the potential endogeneity of unemployment rate, two-stage-least-squares (2SLS) approach is applied thereafter by employing two instruments: Department of Defense annual prime contract awards to each state and state-specific measure of oil price shocks. The effects of both the prime contracts and oil price shocks on state unemployment rate have been well documented by past research.

The results of OLS regression have shown that the effect of unemployment rate is positive and significant at 1 percent level of confidence no matter when the property crimes are taken as one category or each property crime is analysed individually. When it comes to violent crimes, the results are mixed. In the case of estimating the overall violent crimes, the coefficient of unemployment is small and insignificant. As the state-specific linear and quadratic time trends are included one by one, the coefficient of unemployment becomes significant on the 5 percent level of confidence eventually. For the two most serious violent crimes, murder and rape, the effect of unemployment is either significant but wrongly signed or is unstable across specifications. While there is no significant effect of unemployment on assault, there are indeed some evidence for a positive correlation between unemployment and robbery.

The 2SLS estimations have generated similar results: unemployment exerts a consistent, positive and highly significant effect on total property crime. For each individual type, the estimated results are generally supporting a positive correlation between unemployment and each type of property crime, although the coefficient of unemployment is not significant in every specification. For violent crimes, on the other hand, their correlations with unemployment are not quite strong. The coefficient of unemployment is insignificant in each specification for the overall violent crime analysis. Furthermore, the negative correlation between unemployment and murder becomes even more significant in this stage. A similar pattern is also observed for rape. The previously estimated positive correlation between unemployment and robbery becomes unstable across specifications once apply instrument variables.

One interesting finding of this paper is the negative relationship between unemployment and murder, which is contrary to expectation. One possible explanation is that increased unemployment could reduce the interaction between potential offenders and victims.

2.5 DETECTING THE EFFECT OF LABOUR MARKET OPPORTUNITIES

The labour market opportunities, such as legal income and unemployment rate, have been formally incorporated into one's decision of optimal time allocation between legal and illegal activities, as illustrated in Ehrlich (1973). How much time would be spent on both legal and illegal sectors depends on one's relative expected returns from both activities. As such theory has inspired many researchers, it is very natural that plenty of later works have attempted to test their interested part. Given the previous parts of this literature review, some papers are interested in demonstrating the deterrent effect of law enforcement as it is the fundamental assumption of the theory. Meanwhile, other papers have shown their intentions to identify the relationship between crime and unemployment because the effect of unemployment is predicted to be ambiguous. Similarly, there are papers trying to examine the effect of labour market status which include relatively more complete information of legal labour market.

Doyle, *et. al.* (1999) is one of the examples that aim to test the role of labour market on crime. In this paper, the labour market conditions are represented by the average income, unemployment rate and income inequality. While the income level and unemployment rate are familiar to the theoretical framework, income inequality is not incorporated. In order to justify the inclusion of income inequality, Freeman (1996) is used as support given that it has predicted a positive relationship between income inequality and crime. Furthermore, the effects of income level and unemployment rate are far from clear. As analysed previously, unemployment rate could have two counterbalancing effects on crime: opportunity and motivation. Likewise, the income level could capture more than just the potential payoff from legal labour market: it could also measure the potential gain of illegal activities such as property crimes.

In order to test the net effects of labour market components, a set of panel data has employed which covers 48 contiguous U.S. states plus the District of Columbia for the years 1984-1993. The basic model is specified as

$$\begin{aligned}\ln crime_{it} &= \beta_1 \ln wage_{it} + \beta_2 \ln unemployment_{it} + \beta_3 \ln gini_{it} \\ &+ \beta_4 \ln ymen_{it} + \beta_5 \ln pap_{it} + \beta_6 \ln police_{it} + \alpha_i + \varepsilon_{it}\end{aligned}.$$

The dependent variable represents the per capita property crime which is the sum of larceny, burglary and automobile-theft divided by the state population. The independent variables include real annual wage[6], unemployment rate, Gini coefficient,

[6] The real annual wage of all industries enters the equation first. Then the annual wage is separated by industries, such as agricultural services, mining, construction, manufacturing, transportation and public

percentage of young male aged between 15 and 29, probability of arrest and per capita police officers. In addition, the model also includes a state-specific effect to count for the unobservable factors that differ across states.

As law enforcement instruments, arrest rate and per capita police, are potentially endogenous to the system, OLS approach will generate inconsistent estimates. One effective solution to correct such biases will be applying instrument variables. According to relevant literatures, four instruments have been selected to count for the endogeneity of arrest rate and per capita police: arrest rate for violent crime, per capita police in neighbouring states, per capita personal tax revenue and the percentage of the vote cast for the Republican candidate in the biannual elections for U.S. representatives. Given the applying of instrument variables, the methodology of General Method of Moments (GMM) has been carried out.

As further investigation, the same procedure of analysis has also been applied on violent crime rate which is the sum of per capita murder, rape, assault and robbery. The aim of such estimation is to determine whether violent crime is also responsive to labour market conditions. The violent crime model has employed the same set of explanatory variables except one adjustment: the arrest rate of violent crime is instrumented by the arrest rate of property crime along with the other three instruments mentioned earlier. The estimation is also carried out with GMM technique.

---

utilities, wholesale and retail trade, finance, insurance and real estate, services, and government, and re-enter the model.

The estimated results can be summaries by a few points. Firstly, the real annual wage has a substantial negative effect on both property and violent crimes. Once the annual wage is disaggregated by industries, the wage in low-skilled sector of wholesale and retail trade has a negative and significant effect on property crime. This result is consistent with the expectation that the wages of low-skilled sectors would mainly influence the crime rate through the motivation effect because they are measuring the legal income of more crime-prone individuals. Secondly, income inequality is largely insignificant in explaining both property and violent crime. Thirdly, the proportion of young male aged between 15 and 29 has shown positive and significant effect on property crime but negative and significant effect on violent crime.

Gould, *et. al.* (2002) has focused on the relationship between labour market conditions of unskilled men and crime because it argues that unskilled men are most likely to commit crimes and their declining wages and employment opportunities are expected to increase their involvement in crime.

To test this argument, this paper has employed county-level panel data covering the years 1979-1997 for United States. The "core" model takes the index crime rates as dependent variables.[7] The independent variables of interest are the weekly wages for non-college-educated men, unemployment rate of non-college-educated men and per capita income. County-specific and year-specific dummy variables are also included to count for the county-level and yearly unobservable heterogeneity. In addition, each specification also controls for changes in the age, sex and race composition of the county. The estimated results indicate that all three economic variables are very

[7] The index crimes include auto theft, burglary, larceny, aggravated assault, murder, robbery and rape. The estimations have also been applied on the aggregated property crime, violent crime and overall crime.

significant in explaining each index crime except rape. More specifically, each economic variable has the expected sign: the weekly wage of less educated men is negatively correlated with crime rates and the unemployment rate of them has positive effect on crime rates. Furthermore, the per capita income, which is supposed to measure the economic prosperity, constantly has positive and significant effect on crime rates.

Based on the “core” model just introduced, the analysis has been extended by incorporating variables measuring county-level crime deterrence. The newly included variables are the county-level arrest rate, state expenditure per capita on police, and state police employment per capita. In the extended model, the coefficient of less educated wage remains significant for both property and violent crimes, although the magnitudes have dropped. The unemployment of less educated, on the other hand, is still significantly correlated with property crimes but loses significance for violent crimes. However, the sign of unemployment rate is always positive as expected.

Cohen, *et. al.* (1980) also aims to test the effects of labour market conditions on property crimes. In addition, it also predicts a negative relationship between the population density in physical locations and the risk of being property crime victims with other things being equal. Such prediction is resting on the opportunity theory mentioned earlier. It is basically arguing that the occurrence of criminal activities requires the simultaneous interaction of three elements: motivated offenders, suitable targets and the absence of effective guardians. In this theory, the effective guardians are referring to the people able to prevent the violation from occurring either by their physical presence or certain forms of actions. The residential population density is hereby calculated and assumed to measure the strength of guardians of primary-group locations. An increase in the residential population density is expected to deter the occurrence of crimes which mainly take place in or near people's residences.

The relationships of interest between crime and labour market conditions as well as the residential population density are tested by employing time series data for the years 1947-1972. The dependent variables are the crime rates of robbery, burglary and automobile theft for the United States. The independent variables include once-lagged crime rate,[8] residential population density, unemployment rate, percentage of population aged between 15 and 24, total consumer expenditures and par capita automobiles. All the independent variables are measured on national level.

For robbery, all the variables of interest have obtained reasonable signs. The residential population density is negatively correlated with robbery and the estimated coefficient is highly significant. This result has confirmed the predicted guardian effect that residential population could have on crimes. Unemployment rate has obtained a significant and negative relationship with robbery which is not difficult to explain. As claimed in the paper, the main contribution of the unemployment variable is to take into account the effect of business cycle on levels of exposure to the risk of robbery. With higher unemployment rate representing an economic downturn, larger fraction of individuals would have lower probability to expose to motivated robbers. In addition, the percentage of young people has shown positive and significant correlation with robbery as expected.

[8] The coefficient of once-lagged crime rate will tend to be zero if changes in other exogenous variables have fairly rapid impacts on the property crime rates and significantly different from zero if the changes in exogenous variables still affect crime rates after a year.

The estimation for burglary has generated similar results: it has confirmed that burglary rate is negatively correlated with residential population density and positively correlated with young people proportion. Contrary to the case of robbery, unemployment rate has obtained positive and significant coefficient in the estimation for burglary. The explanation could be that unemployment rate has picked up the motivation effect and higher unemployment rate indicates more motivated burglars. The estimation has also found an interesting yet reasonable relationship: the consumption of non-automobile durable goods has positive and significant correlation with burglary rate.

In the case of automobile theft, the residential population density and young people proportion still have the expected signs: negative for residential population density and positive for young people proportion. The coefficients are significant in both cases. The unemployment rate, on the other hand, has switched back to negative signs. This result could indicate that people not working might be less likely to expose their automobiles to illegal removal. As the per capita automobile has entered the estimation to capture the unique feature of automobile theft, no significant relationship has been detected between the two.

2.6 DETECTING THE EFFECT OF YOUNG PEOPLE POPULATION

The percentage of young people has been constantly incorporated in empirical models of crime as an explanatory variable, which can be seen in our previously reviewed literature. Although demographic variables do not take parts in the classic theories of crime, they have been customarily included in empirical estimations helping to explain the variations in crime. The choices of demographic variables depend on the specific situation of each country. For example, it seems necessary to include both percentage of black people and percentage of young people for the U.S. literatures. However, the former would seem excessive for the works in the U.K. given its ethnic composition. The percentage of young people, on the other hand, is more universal across cases as it has been broadly accepted that there exist a robust relationship between age and criminal involvement. As stated in Levitt (1999), “there is a sharp rise in criminal involvement with the onset of adolescence followed by a steady decline with age. The prime ages for criminal involvement are roughly 15 -24. Property crime typically peaks somewhat earlier than violent crime.” According to such statement, one would be reasonable to predict that as the fraction of the population most prone to involve in crime rises, aggregate crime is likely to rise.

Searching for the reason, some would attribute the high propensity to engage in crime of young people to their labour market situations. Essentially, young people are expected to respond to incentives and opportunities the same way as adults do. The difference is that young people would expect lower returns from legal labour market given their lack of experiences and qualifications. The relatively lower opportunity cost from legal activities would therefore lower young people’s desired payoffs from illegal activities and increase their propensity to involve in crime. Furthermore, there are evidences showing that offenders under the age of 18 would be treated with much milder punishment once convicted and, in addition, their criminal records would be sealed after the age of 18 so that they would not affect the offenders’ future career prospects. All these facts indicate a lower expected cost of criminal involvement for young people who can become more “carefree” when deciding whether or not to commit crimes.

Levitt (1999) has designed a simple approach to test the impact of changing age distribution on aggregate crime rates. As pointed out in the paper, one plausible strategy of testing such impact is to run reduced-form regression with the aggregated crime rate being the dependent variable and the age distribution of population and other control variables being the regressors. Such approach has been adopted by numerous literatures. While some of them have found significant relationship between crime rate and young people proportion, such as Cohen and Land (1987), majority of relevant works have failed to uncover a significant effect between the two. Alternatively, Levitt (1999) has applied another approach which decomposes the crime rate by ages. Then by taking the age-specific crime rate in a particular year as given, the hypothetical aggregate crime rate can be computed using the age structure from a different point in time. For example, between 1960 and 1980, the percentage of population aged between 15 and 24 has risen from 13.4 percent to 18.7 percent in the United States. The following 15 years, 1980-1995, have almost completely undone this rise: it has dropped from 18.7 percent to 13.7 percent. Using the proposed approach, it is possible to calculate how much in the changes of crime rate can be explained by the changes in age distribution.

The paper has investigated the changes in three types of crime: murder, violent crime, and property crime. For the period 1960-1980, the murder rate rose from 5.08 to 10.22 per 100,000 population indicating an increase of more than 100 percent. Changes in the age structure are estimated to count for one-fifth of that total rise. For both violent and property crime, changes in the age distribution contribute to similar rises, 17 and 22 percent respectively. For the period between 1980 and 1995, the changing age distribution has lowered crime rates due to the declined percentage of population aged 15 to 24. For example, 40 percent of the decrease in murder over the period can be explained by the changing age structure. Furthermore, the benefit of aging population can count for 12 percent and 18 percent declines in violent and property crime respectively.

Aside from the worse labour market prospects of young people, the expected milder punishment can partially explain their higher propensity to engage in crime, as analysed previously. Levitt (1998) has constructed an empirical model to test this proposition. The basic empirical model takes the number of juvenile crimes per juvenile aged 15-17 as dependent variable. The variations of dependent variable are explained by once-lagged juvenile custody rate[9], which measures the severity of juvenile punishment, along with other controlling variables. Such variables include the percentage of black people, the percentage of metropolitan residents, unemployment rate, legal drinking age, and the age distribution of population. As the empirical model is applied on state-level panel data over the period 1978-1993 in the United States, both state-specific and year-specific dummy variables have been included to control for unobservable heterogeneity. The estimated results have shown that juvenile crime is responsive to harsher punishment. As the same empirical analysis has also been applied on adult crime rate, there is evidence to claim that the estimated decrease in juvenile crime rate caused by increased custody rate is at least as large as the corresponding reduction in adult crime rate due to a same rise in

[9] The juvenile custody rate is measured by the number of juvenile in custody per juvenile aged between 15 and 17.

custody rate. In addition, there are sharp changes in the crime rates associated with the transition from the juvenile to adult court. As soon as turning to the age of majority, states with harsher adult punishment relative to juvenile punishment see sharper drops in crime rates comparing to states with milder relative adult punishment.

Chapter Three: Time Series Analyses

3.1 INTRODUCTION

In this chapter, we aim to identify the relationship between unemployment and crime in England and Wales by adopting time series analyses. There are several factors that motivate us to do this work. First of all, the unemployment – crime relationship has been one of the focal points in the economic literature of crime. The relationship of interest, however, is still far from clear. On the one hand, the effect of unemployment and crime is predicted to be ambiguous by Ehrlich (1973) and Cantor and Land (1985). Ehrlich argues that, while an increase in unemployment rate will increase people's participation in illegal activities through reducing their opportunity of doing so, higher unemployment will also increase one's demand for wealth due to his higher probability of ending up with the least desired situation – unemployed in legal sector and failed in illegal activities, and thus reduce his incentives to commit crimes.[10] Cantor and Land claim that increased unemployment could have two offsetting effects on crime: reducing criminal opportunities and, meanwhile, motivating potential offenders to commit crimes.[11] On the other hand, empirical studies testing the unemployment – crime relationship have obtained mixed results. For example, Reilly and Witt (1996), Witt *et al.* (1998, 1999), Raphael and Winter-Ebmer (2001) have found positive and significant effect of unemployment on crime rate; Greenberg (2001), Doyle *et al.* (1999) and Entorf and Spengler (2000) found negative and even insignificant U – C relationship. Their results suggest that unemployment has probably picked up both opportunity and motivation effects it could have on crime;

[10] More details of Ehrlich (1973) can be found in Chapter 2.
[11] For more details of Cantor and Land (1985), please see Chapter 2.

and the magnitudes of both effects, and thus the net effect of unemployment, dependent on the specific features of the crime being examined as well as the empirical model estimated.

Given the above facts, we argue that it could be useful to implement time series analysis because, despite of the co-existence of opportunity and motivation effects, the unemployment – crime relationship could be stable in long-run. Applying time series analysis enables us to explore whether there is an equilibrium correlation between unemployment and crime. Furthermore, unlike panel data and cross-sectional analyses, time series analysis only depends on the variations in variables over time by diminishing the spatial deviations. Therefore, it is possible that applying time series data and techniques could generate different results.

The second motivation is that the literature in England and Wales being surveyed in this chapter apply their time series analyses by applying different data set as well as different variables and, in some cases, reach different conclusion. For example, Hale and Sabbagh (1991) investigate the unemployment – crime relationship based on national level time series data covering the period 1949-1987. They find no long-term cointegration between unemployment and crime. Pyle and Deadman (1994a) examine the period 1946-1991 searching for long-term correlation between property crimes and business cycle. Their analyses suggest that the changes in property crimes are all cointegrated with unemployment in equilibrium. As seen in these two papers, as well as those studies in our literature review section, changing the time period examined and the variables incorporated could generate different results for the unemployment – crime relationship. And this is particularly true for time series analysis because all the information used in the analysis comes from the variations of variables over time. Therefore, we examine a different time period, 1971-2000, and employ a different set of explanatory variables to test the unemployment – crime relationship in our analyses.

In this chapter, we acknowledge the co-existence of opportunity and motivation effects of unemployment and focus our interest on testing both long-term and short-term relationship between unemployment and crime using co-integration and error correction techniques. We employ national time series data covering the years 1971-2000 in England and Wales and carry out analyses on overall and individual property crime rate including burglary, theft and handling, and fraud and forgery. (For overall crime and fraud and forgery, we only include the years 1971-1997 to avoid the influence of new counting rules adopted in 1998, which will be given detailed description later.) We only include three explanatory variables, including crime-specific detection rate, custody rate and unemployment, to maximally utilize from our sample size of 30. In addition to unemployment, which is our main concern, we include detection and custody rates in order to eliminate their effects on crime because they should most directly affect crime and cannot be omitted from the specification.

We have mainly found that each crime rate being tested has cointegration relationship with the three explanatory variables. Detection rate is negatively cointegrated with the overall crime and positively cointegrated with property crime rates in long-run; custody rate constantly has negative long-term correlation with all the crime rates; and unemployment has positive cointegration with the overall crime as well as burglary and theft, while it has negative cointegration with fraud. In short-run, the change of each crime rate is affected by contemporary changes in explanatory variables, but not affected by their lagged changes.

The chapter is structured as follows. Section two reviews relevant literature from both theoretical and empirical aspects. In section three, we describe the employed data and present summary statistics. Section four explains the empirical models and the estimation methodologies, while section five reports the results as well as their interpretations. Section six briefly summarizes the main findings of this chapter. It will also discuss the limitations of this work as well as potential future improvement.

3.2 LITERATURE REVIEW

Most of the existing empirical studies of crime are based on the theoretical models constructed in Becker (1968) and Ehrlich (1973), and the work in this chapter is no exception. Therefore, this literature review recalls the essential points made by both papers briefly as they have already been introduced in Chapter Two. It then discusses the empirical papers that detecting the relationship between unemployment and crime using time series analysis.

3.2.1 Economic Theory of Crime

Becker (1968) constructs its model by assuming the potential offenders are economically rational and aiming to maximize their expected utility from committing crimes. Therefore, the number of offences one would commit should be affected by his probability of apprehension and the severity of punishment. Specifically, an increase in either the probability of apprehension or the severity of punishment is expected to reduce one's incentives to commit crimes because the increased expected punishment offsets his expected returns from doing so. The aggregated supply of offences can then be derived by assuming all the individuals have the same reaction to the tougher law enforcement efforts and is predicted to be reduced by either higher probability of apprehension or more severe punishment, or both.

The model in Ehrlich (1973) is developed by allowing each individual to freely allocate his time between committing crimes and working in legal sectors. His aim, however, is still maximizing the expected utility. Other things being equal, an increase in either the probability of arrest or the severity of punishment would reduce one's participation in crimes by reducing the relative returns between illegal and legal activities. Similarly, better opportunities from legal labour market will also decrease the number of offences one would commit by making the legal sectors more profitable comparing to committing crimes. The unemployment rate enters one's expected utility function measuring the risk of legal labour market. On the one hand, an increase in the unemployment rate is expected to increase one's participation in crimes due to his lower opportunity cost of doing so. On the other hand, however, higher unemployment rate will increase one's probability to end up with the lease desired situation – failed in illegal activities and unemployed in legal sectors. Such change will diminish his willingness to take the risk of committing crimes. Therefore, as argued in Ehrlich (1973), the net effect of higher unemployment rate on crime is ambiguous to predict.

According to the theories introduced above, we carry out time series analysis to test the relationships specified as follows:

$$C = \Psi(P, F, U),$$

where the aggregated crime rate C is a function of the probability of apprehension P, the severity of punishment F, and the unemployment rate U. we use the detection rate and the number of people in custody per 1000 population as proxies for probability of apprehension and severity of punishment. Thus, we expect both variables to have negative correlations with crime rates. Meanwhile, we accept the possibility that the effect of unemployment on crime is not necessarily positive and significant and wait to see what the analysis can reveal.

3.2.2 Empirical Studies

The early time series analyses of crime do not benefit from the cointegration and error correction techniques and are mainly based on ordinary least squares (OLS) regressions (e.g. Cantor and Land 1985; Britt 1994). These papers are criticized by later studies for not considering the stationarity of time series variables. Because, as they argue, simply applying the OLS estimation between non-stationary variables will lead to spurious and invalid results due to their different trends over time. Therefore, with the development of cointegration technique, more recent time series papers have been able to avoid the spurious results when analysing non-stationary variables by adopting such approach. This section will mainly focus on the literature investigating the relationship between unemployment and crime in England and Wales by using cointegration and relevant analysis.

Cantor and Land (1985) make significant contribution to uncover the unemployment – crime relationship by suggesting that an increase in unemployment could have double impacts on crime: opportunity and motivation. While higher unemployment rate could reduce the opportunities for certain types of crime, it will also motivate more people to participate in criminal activities by lowering their opportunity cost of doing so. In order to test this theory, they have designed an empirical model by assuming the change in crime rate is affected by two variables: the contemporary and first-differenced unemployment rates. While the former is expected to capture the opportunity effect and thus have negative effect on crime, the latter is assumed to measure the motivation effect and should have positive effect. By applying time series data in the United States covering the period 1946-1982, they have indeed found that increased contemporary unemployment rate broadly exhibits negative effect on both violent and property crimes, suggesting a significant opportunity effect it is supposed to capture. On the other hand, an increase in the first-differenced unemployment rate only shows positive correlation with property crimes, indicating that increased unemployment rate only motivates the potential offenders looking for financial benefit.

Britt (1994) adopts the strategy proposed in Cantor and Land (1985) to test the unemployment – crime relationship for young people in the U.S. over the years 1958-1990.[12] By using the crime-specific arrest rates as proxies for different types of crime rates, he has obtained very similar results to those of Cantor and Land (1985). For both violent and property crimes, an increase in the contemporary youth unemployment is negatively associated with the annual changes in the youth arrest rates.[13] Meanwhile, the first-differenced youth unemployment only positively affects

[12] The young people are defined as the persons aged 16 to 19 years old and both unemployment rate and arrest rates (used as proxies for crime rates) are restricted to this age group.
[13] The violent crimes include homicide, rape and aggravated assault. The property crimes refer to robbery, burglary and larceny.

the annual changes in youth property arrest rates, but has no significant impact on violent offences of young people.

Greenberg (2001) and Hale and Sabbahg (1991) criticize the methodology adopted in Cantor and Land (1985) for ignoring the fact that variables could be different in their orders of integration. They argue that, when the variables are non-stationary (i.e. having an order of integration higher than zero) and integrated at different orders, the classic OLS estimation breaks down and special analysing procedures are needed. Accordingly, Greenberg (2001) applies cointegration and error correction techniques to test the relationship between divorce rate, unemployment rate and the crime rates of homicide and robbery. Based on annual time series data in the U.S. over the years 1946-1997, both homicide and burglary have been found to be positively cointegrated with divorce rate. Furthermore, in short-run, the change in homicide rate is positively affected by the change of divorce rate, and negatively correlated with one year lagged unemployment rate. Meanwhile, the error correction model (testing the short-run relationship) has found similar result for robbery: the change in robbery is positively correlated with the change in divorce rate and negatively correlated with that of once-lagged unemployment rate.

Hale and Sabbagh (1991) examines the unemployment – crime relationship in England and Wales using annual time series data covering the period 1949-1987. Having decided that unemployment and crime rates[14] are all I(1) series (the variables are non-stationary on their levels but stationary after first differencing), he applies Engle-Granger two step procedure to detect for long-term cointegrating relationships

[14] The crime rates being tested are total theft, theft by an employee, shoplifting, handling stolen goods, auto theft, total burglary, robbery, and violent crime.

between unemployment and crime. As the results generate no cointegration between the concerned variables, however, he specifies an alternative empirical model where the change in crime rate is correlated with both contemporary and once-lagged change in unemployment. Consequently, the estimation results show that only the crime of theft by an employee is negatively affected by the contemporary change in unemployment, suggesting an opportunity effect. Meanwhile, the crime rates of theft, burglary, and robbery have all shown positive correlation with the current change of unemployment, which is against his approach of using first-differenced unemployment rate to capture the opportunity effect.

Pyle and Deadman (1994a) is interested in testing the relationship between business cycle and property crimes using both annual and quarterly time series data in England and Wales. While the crime rates being examined are theft and handling stolen goods, burglary and robbery, the explanatory variables include conviction rates, the number of police officers, the number of males aged 15-19, unemployment rate, real personal consumption, GDP, and a weather index. By finding that the crime rates are I(2) series (need to be differenced twice for stationary) while the explanatory variables are I(1) series (need to be differenced once to become stationary), they try to look for cointegrations between the first-differenced crime rates and the explanatory variables. Based on the annual time series data covering 1946-1991, the first-differenced crime rates have all been found to be cointegrated with the economic variables. Specifically, while an increase in either GDP or personal consumption is negatively associated with the changes of all property crimes; higher unemployment rate has positive effect on them. Furthermore, the conviction rate is found to have negative cointegration with the growth of robbery and theft over the examined period. Further analyses using quarterly data, which is available for 1975(1)-1991(4), have confirmed the previous finding: GDP is negatively correlated with the growth of all the property crimes in both long-run and short-run models.

Based on the error correction model estimated in Pyle and Deadman (1994a), Deanman and Pyle (1997) try to forecast the levels of property crimes for the years 1992-1996. The aim is to see how well the model estimated in the previous paper performs as a forecasting device. As they adopt a strategy that forecasts the one year ahead growth in crime rates (i.e. from one year to the next), the predicted values suggest that the error correction model in the previous paper works quite satisfactorily in forecasting the trends in crime rates. Particularly, the forecasted crime rates have picked up the 1992-turning point for theft and the 1993-turning point for burglary. However, the model works less accurately in predicting crime levels and it tends to "exaggerate" the actual values for the examined period.

Hale (1998) criticizes Pyle and Deadman (1994a; 1997) on the foundation of their model specification: all the crime rates being tested are integrated of order 2. By re-examining the period 1946-1991, Hale shows that the crime rates of theft and handling, burglar and robbery are all I(1) series, instead of I(2). This result suggests that both cointegration and error correction models specified in Pyle and Deadman (1994) are wrong. Furthermore, as Hale re-estimates the cointegrations according to the correct specification, burglary and theft are found to be cointegrated with only personal consumption while robbery has no cointegrating relationship with any of the explanatory variable. These long-term relationships are less strong than what is found in Pyle and Deadman (1994a) and such difference should be due to the miss-specified equation, as argued in Hale (1998). In the short-run dynamic models, the changes in burglary and theft are all positively affected by the change of unemployment, negatively correlated with the change in personal consumption, the number of police officer, as well as the conviction rate.

Dhiri *et al.* (1999) carry out time series analyses based on the period 1951-1998 in order to present a projection of property crimes in England and Wales for the following three years: 1999-2001. The Engle-Granger two step procedure reveals that theft and burglary both have cointegrating correlations with two explanatory variables: the stock of crime opportunities (measured by the sum of personal consumption for the past 3 years) and the number of young males. Having established cointegrations between crime rates and the concerned explanatory variables, they go on to estimate the dynamic models using the approach developed in Sims *et. al.* (1990). The results show that, in short-run, the growth in theft and burglary are all positively affected by the one year-lagged growth in the number of young males as well as the once-lagged change in the stock of crime opportunities.

Most of the literature introduced previously use Engle-Granger two step procedure to test for cointegrations between crime rates and explanatory variables. An alternative approach to test for cointegration is Johansen technique. Saridakis (2008) adopts such methodology trying to establish long-run correlations between violent crimes and the chosen control variables. The aim of this work is to examine the hypothesis that tougher punishment and better economic opportunities will reduce violent crime in equilibrium using time series data in England and Wales during 1960-2000. While the dependent variables being tested are the overall violent crime, rape, indecent assault on a female and aggravated assault, the independent variables include conviction rate, imprisonment rate, male unemployment rate, poverty rate, and beer consumption. The results from applying Johansen technique suggest that cointegration relationship does exist for aggravated assault. Furthermore, the estimated cointegrating vector strongly supports the expected effect of law enforcement instruments: both conviction and imprisonment rates are negatively correlated with aggravated assault. They also find that the crime rate of aggravated assault is, in long-run, positively affected by male unemployment rate, poverty rate, as well as beer consumption. Meanwhile, for the more serious violent crimes such as rape and indecent assault on female, there is no long-run correlation detected between them and the concerned explanatory variables.

Time series data is one of the most important and commonly used data structures to investigate the determinants of crime, with the other structures being cross-sectional and panel data.[15] However, Levitt (2001) argues that "national-level time series data are an extremely crude tool for answering criminological questions" for several reasons. Firstly, although time series data is an ideal tool for analysing macro variables such as economic growth and inflation etc., crime rates and their influential variables usually exhibit significant local variations. Thus, applying time series data on crime modelling is not able to pick up the extra information that is varying by locations. Secondly, time series data usually provide limited sample size comparing to panel and individual data. Hence, it is difficult to include a wide range of explanatory variables into the equation. Thus, the estimated parameters only reflect the correlations between explained and explanatory variables, instead of causal links. This is because, in order to interpret the coefficients as causal, it is necessary to include all

[15] Individual-level data are less broadly employed due to the unavailability.

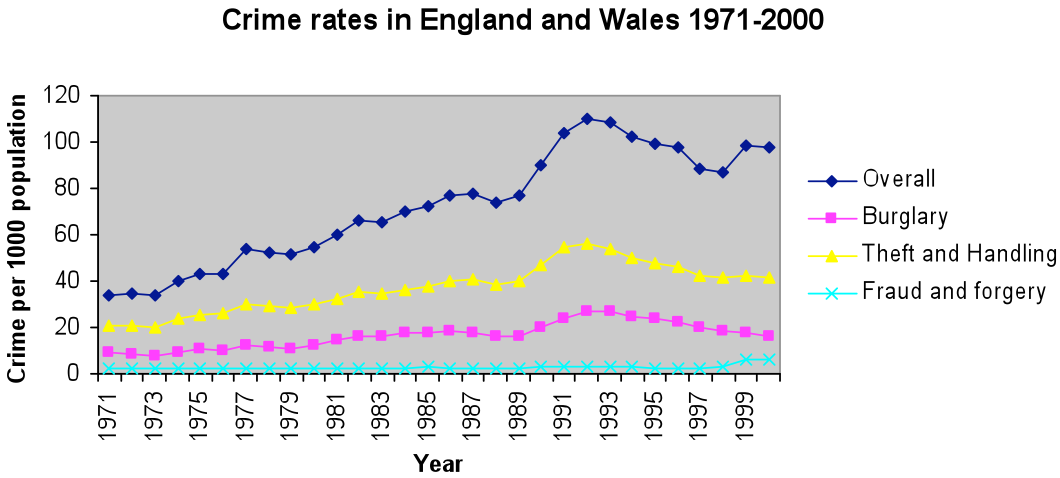
the potential crime-influential factors into the equation. With the limited degree of freedom given by time series data, it is apparently not rational to do so. Thirdly, some papers are particularly interested in separately estimating the two channels through which unemployment affects crime: opportunity and motivation, as predicted in Cantor and Land (1985). Using national-level time series data, however, has limited power for such job.

Given the potential shortcomings of time series data as mentioned above, we argue that applying time series data and techniques is still meaningful for investigating the correlation between crime and its relevant factors. In this chapter, we are particularly interested in using national-level time series data in England and Wales to test the unemployment – crime relationship, as well as the deterrent effect of law enforcement. Our reasons are as follows. Firstly, given the co-existence of opportunity and motivation effects of unemployment on crime, the net effect of unemployment could be stable over time. Therefore, by carrying out cointegration analysis, we can reveal that, in long-run, whether the net effect of unemployment is positive, negative or insignificant. Secondly, we have applied both panel and spatial analyses in England and Wales to identify the relationships between crime rates and unemployment, as well as other relevant variables. Using time series data and techniques could provide a different angle to investigate this issue and perhaps generate interesting results. Therefore, in this chapter, we examine the correlation between crime rates and unemployment rate, along with law enforcement instruments by employing national level time series data in England and Wales.

3.3 DATA DESCRIPTION

In this section, we introduce the basic properties of both dependent and independent variables over the examined period 1971-2000. The explained variables are the overall crime rate as well as individual property crime rates: burglary, theft and handling, and fraud and forgery. We mainly choose property crimes to analyse for the following reasons. First, from the theoretical aspect, we expect the property crimes to be more responsive to the included explanatory variables than violent crimes. The economic models of crime in Becker (1968) and Ehrlich (1973) suggest that relevant variables affect crime rates through changing the expected costs and payoffs from committing crimes. Therefore, an increase in the expected punishment will deter property crimes by increasing the expected costs of illegal activities, while higher unemployment rate could encourage the involvement in property crimes through reducing the opportunity cost of doing so. The violent crimes, on the other hand, show less direct correlations with these factors because the targets of such crimes are not financial benefits. Second, the previous empirical analyses tell us that property crimes are much better explained by the economic models than violent crimes. In both chapters of panel data and spatial analysis, the law enforcement instruments constantly have negative and significant effects on property crimes, while they show less consistent correlations with violent crimes. Meanwhile, the social-economic factors exhibit significant (or insignificant but explainable) effects on property crimes while they have, in general, shown no systematic correlation with violent crimes. Furthermore, Saridakis (2008) has demonstrated that only minor violent crime, such as aggravated assault, is correlated with law enforcement variables as well as social-economic factors. More severe crimes, such as rape, indecent assault on female, are not affected by these variables in long-run. Based on the reasons mentioned above, we only conduct our analysis on the overall and individual property crimes.

The crime rates are measured by the number of offences recorded by police per 1000 population. The data covering the period 1971-2000 are obtained from the Home Office publication *Criminal Statistics*. The following graph shows the time trends of the four crime rates over the examined period.



As seen in the chart above, burglary and theft have quite similar shapes over time. They both keep increasing since the year 1971 and peak around the years 1992-1993. After that, they both experience mild and stable reduction until the year 2000. The line at the bottom of the chart indicates the time trend for fraud and forgery. It has much lower crime rate, which is less than 10 offences per 1000 population during 1971-2000, comparing to burglary and theft. Its trend is slightly increasing until a sharp jump in 1998, which is due to the introduction of new counting rules discussed in previous chapters. The overall crime rate is lying on the top of the chart and has clearly picked up the trends of burglary and theft before the year 1998. The reason is obvious: while theft and handling takes more than 50 percent of the total crime, burglary takes around 30 percent. Thus, the shape of total crime rate is unsurprisingly dominated by the movements of theft and burglary. However, during 1998-2000, while both theft and burglary keep decreasing, the overall crime rate has an obvious upward jump because of the new counting rules. Therefore, in order to avoid the structure breaks, we have deleted the years 1998-2000 from the sample when empirically analysing the overall crime rate as well as fraud and forgery.

Table 3-1 below reports the summary statistics for each crime rate being tested. The number of observation shows that, if no value is missing, the sample size is 30 in later analyses.

Table 3-1

Overall and property crime rates

		Property crimes		
	Overall crime	Burglary	Theft and handling	Fraud and forgery
Mean	71.98	16.53	37.14	2.85
Median	73.04	16.38	38.11	2.60
Maximum	109.64	26.86	56.05	6.20
Minimum	33.69	7.81	20.31	2.04
Std. Dev.	24.24	5.49	10.25	0.97
Observations	30	30	30	30

The crime-specific detection rate is used as a proxy for the probability of apprehension and expected to be negatively correlated with crime rate. As explained in previous chapters, the detection rate measures the percentage of recorded offences that have been solved by the police through giving caution, fine or charge. The data is collected from *Criminal Statistics*. Meanwhile, we use the number of offenders sentenced into custody per 1000 population as a proxy of severity of punishment. According to the theories of crime, this variable is expected to be negatively correlated with crime rates due to its deterrence and incapacitation effects. The data source is the website of National Statistics. We use the unemployment rate defined by the International Labour Organisation (ILO), instead of the claimant count rate[16], in the time series analysis and the data is also from the website of National Statistics. The ILO unemployment rate is referring to the number of people who are looking for and available for work as a proportion of the resident economically active population. The correlation between this variable and crime rates are ambiguous to predict because of the opportunity and motivation effects it has.

Table 3-2 and 3-3 below summarize the basic statistics for the explanatory variables.

Table 3-2
Overall and property detection rates

		Property crimes		
	Overall crime	Burglary	Theft and handling	Fraud and forgery
Mean	34.83	27.07	32.60	64.93
Median	35.00	28.00	34.50	68.50
Maximum	47.00	37.00	43.00	84.00
Minimum	24.00	12.00	17.00	29.00
Std. Dev.	7.29	6.58	8.18	16.36
Observations	30	30	30	30

Table 3-3
Other independent variables

	People in custody per 1000 population	ILO unemployment rate
Mean	0.93	7.73
Median	0.88	7.35
Maximum	1.26	12.1
Minimum	0.75	3.60
Std. Dev.	0.14	2.68
Observations	30	30

3.4 EMPIRICAL MODELS AND METHODOLOGIES

One of the advantages of using time series data and technique is that we are able to model both long-run and short-run relationships among variables. In the presence of cointegration, we can specify an error correction dynamic model (ECM) accordingly

[16] The claimant count rate is used in the panel data and spatial analysis chapters because it is available on local authority level.

and estimate the short-term relationship between the differenced variables. Following this procedure, we firstly define our long-term cointegration model as

$$\ln(crime)_t = \beta_0 + \beta_1 \ln(\det ection)_t + \beta_2 \ln(custody)_t + \beta_3 \ln(ilo)_t + \beta_4 trend\ , \quad (3.1)$$

where the crime rate in period t is assumed to have equilibrium relationship with contemporary detection rate, number of people in custody, ILO unemployment rate as well as a constant and linear time trend. As analysed previously, while the coefficient of unemployment is difficult to predict, both detection rate and people in custody are supposed to have negative correlation with crime rate.

To test the cointegration given by equation (3.1), we follow the Engle-Granger two step procedure developed in Engle and Granger (1987). However, in performing such examination, the first step is to determine the order of integration for each variable because, by definition, cointegration necessitates that the variables included in the equilibrium function should be integrated of the same order. Therefore, we start by applying both augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests on the level of each variable to test for unit root.[17] For instance, the level of crime rate is tested by equations (3.2) and (3.3) which respectively indicate ADF and PP tests:

$$\Delta C_t = \alpha_0 + \gamma C_{t-1} + \alpha_1 t + \sum_{i=1}^{p} \theta_i \Delta C_{t-i} + u_t \quad (3.2)$$

and

$$\Delta C_{t-1} = \alpha_0 + \gamma C_{t-1} + \alpha_1 t + u_t\ . \quad (3.3)$$

In both tests as shown in equations (3.2) and (3.3), our primary concern is the coefficient of once-lagged crime rate, γ. In the case of $\gamma = 0$, the crime rate C

[17] See Dickey and Fuller (1979) and Phillips-Perron (1988).

follows a random-walk process indicating crime rate on level is non-stationary. On the other hand, if $\gamma \neq 0$, the crime rate C cannot be written in the form $C_t = C_{t-1} + u_t$. In such case, we can conclude that there is no unit root in the crime rate and thus it is stationary. As seen in equation (3.2) and (3.3), the ADF test includes certain number of lagged dependent variable to eliminate the serial correlation in the residuals because the distribution theory supporting the ADF test is based on the assumption that the residuals are uncorrelated and have a constant variance. The PP test, on the other hand, is a generalised ADF test and requires milder assumptions on the residuals. Instead of incorporating lagged dependent variables to the right-hand side of the test equation, it avoids the bias of correlated residuals by calculating the correct *t* statistic for the unit root coefficient. Furthermore, we use the critical values derived according to MacKinnon (1996) for both ADF and PP tests to decide whether we should reject the null hypothesis that a unit root exists (the variable being tested is non-stationary).

In both unit root tests, a constant is always included when testing the level of crime rate. However, whether a time trend is incorporated depends on its movement: the time trend is included when the crime rate has either increasing or decreasing tendency. If a unit root is found (i.e. $\gamma = 0$) based on these tests, we cannot reject the null hypothesis that the crime rate is non-stationary and, in such case, we take the first-difference and apply the unit root tests again.[18] We repeat this process until the differenced crime rate becomes stationary (i.e. $\gamma \neq 0$) and the times crime rates has been differenced is his order of integration. Furthermore, we apply the same procedure of unit root test on all the explanatory variables as well. [19]

[18] We include neither a constant nor a time trend when testing the unit root for a differenced variable.
[19] The results of unit root tests are reported in the next section.

Under the condition of all the variables have the same order of integration, we keep on our analyses by performing the Engle-Granger two step procedure to test for cointegration. With OLS regression, we first of all estimate the long-run equilibrium function given by equation (1) and obtain the estimated residuals. The next step is to check the residuals for the order of integration. If they are stationary, we can conclude that there is cointegration among the included variables and the OLS regression yields "super-consistent" cointegrating parameters. On the other hand, in the case of the residuals are non-stationary, we have to reject the null hypothesis that the variables are cointegrated in equilibrium and the OLS estimation generates spurious results.[20]

Based on cointegration relationship, we hereby specify an error correction mechanism (ECM) to model the short-term dynamic correlations between the changes in the variables. Formally, we estimate our ECM model represented below using OLS regression for each crime rate:

$$\begin{aligned}\Delta \ln(crime)_t &= \alpha_0 + \sum_{i=1}^{2} \alpha_i \Delta \ln(crime)_{t-i} + \sum_{i=0}^{2} \delta_i \Delta \ln(\det ection)_{t-i} \\ &+ \sum_{i=0}^{2} \gamma_i \Delta \ln(custody)_{t-i} + \sum_{i=0}^{2} \theta_i \Delta \ln(ilo)_{t-i} + \lambda ECM_{t-1}\end{aligned}, \tag{3.4}$$

where the term ECM_{t-1} is the one-year-lagged residuals from estimating the equilibrium equation and calculated as

$$\begin{aligned}ECM_{t-1} &= \ln(crime)_{t-1} - \beta_0 - \beta_1 \ln(\det ection)_{t-1} - \beta_2 \ln(custody)_{t-1} \\ &- \beta_3 \ln(ilo)_{t-1} - \beta_4 trend\end{aligned}. \tag{3.5}$$

While estimating the dynamic model defined by equation (3.4), there are a few things need to point out. Firstly, the error correction model has the advantage of being able

[20] While applying ADF and PP tests to decide the order of integration of the residuals, we do not include either a constant or a time trend in the specification.

to incorporate equilibrium information into a short-run model. The coefficients of first-differenced variables on the right-hand side of equation (3.4) measure the immediate impacts that the changes in detection rate, people in custody and unemployment rate will have on the change of crime rate. Hence, they reflect short-term correlations. On the other hand, the coefficient of ECM_{t-1} is the error correction mechanism and shows how much a drift from equilibrium will be corrected. This coefficient should be negative so that a crime rate above equilibrium level will be pulled back in the next time period to the equilibrium. In this way, the crime rate will maintain a stable long-run relationship with its explanatory variables.

Secondly, we have chosen a maximum lag length of two for the dynamic model. Theoretically, the appropriate lag length should be able to generate the desired properties for the residuals from equation (3.4) such as normal distribution, no serial correlation and no heteroskedasticity. However, we choose to include a maximum of two lags for two reasons: 1) our sample size of 30 does not provide the potential to include a large number of lagged variables. Therefore, we are quite restricted when choosing the explanatory variables as well as the lag length; and 2) as we are using annual data, including two lags is normally adequate to incorporate the dynamic impacts of lagged changes in explanatory variables. In addition, as we have tried to include more lags, the results suggest that including two lags in the dynamic model is long enough to produce the desired properties for the residuals.

3.5 RESULTS

In this part, we will present and discuss our estimation results including those from unit root tests, cointegration tests, as well as dynamic ECM regressions.

3.5.1 Unit Root Test

To determine the order of integration, we apply both ADF and PP tests on each variable. While testing the variables on their levels, we always include a constant. Meanwhile, we decide to include a time trend if the variable constantly rising or falling over the examined period. Our results in table 3-4, 3-5, and 3-6 show that, using both ADF and PP tests, all the variables (excluding the overall detection rate) have obtained insignificant *t* statistics suggesting that we cannot reject the existence of a unit root and they are non-stationary on levels. Therefore, we take the first-differences for the variables (except the overall detection rate) and test for unit root again without including either constant or time trend. Our results indicate that all the first-differenced variables (except the overall detection rate) are stationary and, furthermore, the associated *t* statistics in both ADF and PP tests are significant at 1% level (i.e. the null hypothesis that there is a unit root can be rejected).

Now we move onto discussing the overall detection rate. This variable is stationary on its level according to the ADF test, as shown in table 3-5. However, the PP test shows a unit root for its level according to the insignificant *t* statistic. Given that the ADF test is known as being easier to reject the hypothesized existence of unit root, we decide to take the level of overall detection rate as non-stationary and compute the first-difference for it. As seen in table 3-5, both ADF and PP tests show that the differenced detection rate is stationary based on highly significant (at 1% level) *t* statistics.

Table 3-4
Unit root tests for crime rates

	ADF		PP	
Variable	Level	First Difference	Level	First Difference
Overall	-0.64	-3.26***	-0.85	-3.26***

Burglary	-0.94	-3.64***	-0.20	-3.63***
Theft and handling	-1.29	-3.50***	-0.69	-3.53***
Fraud and forgery	-2.60	-3.86***	-2.05	-3.78***

*** significant at 1% level; ** significant at 5% level; and * significant at 10% level.

Table 3-5
Unit root tests for detection rates

Variable	ADF Level	ADF First Difference	PP Level	PP First Difference
Overall	-3.74**	-3.25***	-2.57	-3.25***
Burglary	-2.96	-4.45***	-1.76	-4.45***
Theft and handling	-0.80	-3.67***	-1.38	-3.68***
Fraud and forgery	-2.66	-2.91***	-2.14	-2.91***

*** significant at 1% level; ** significant at 5% level; and * significant at 10% level.

Table 3-6
Unit root tests for other variables

Variable	ADF Level	ADF First Difference	PP Level	PP First Difference
People in custody	-2.68	-2.97***	-1.83	-2.98***
ILO unemployment rate	-1.42	-3.06***	-0.70	-2.85***

*** significant at 1% level; ** significant at 5% level; and * significant at 10% level.

3.5.2 Cointegration Test

In last section, we have confirmed that all the variables, both dependent and independent, are integrated of the same order, I(1). Therefore, we are reasonable to perform the cointegration tests among these variables following Engle-Granger two-step procedure. For each crime rate, we firstly estimate the equilibrium equation defined by equation (3.1) using the OLS regression and acquire the residuals. Afterwards, we test the residuals with ADF and PP tests to decide whether they are stationary, because if they are, we can accept the variables in equation (3.1) as cointegrated. It is worth noting that neither a constant nor a time trend is incorporated in the unit root tests because the residuals are assumed to have zero mean and constant variances. (Furthermore, we do not find either constant or trend existing in the estimated residuals when checking their scatter plots.) As reported in table 3-7, the residuals from the estimation of each crime rate is stationary on the levels according to both ADF and PP tests with highly significant (at 1% level) *t* statistics. Consequently, we are able to conclude that, although the variables are individually non-stationary over time, their linear combination specified by equation (3.1) is stable over time, i.e. these variables are cointegrated.

The cointegrating correlations are reported in table 3-8 and we summarise the results as following. Firstly, the overall detection rate is negatively cointegrated with the overall crime rate suggesting that, in long-run, higher detection rate comes along with decreased crime rate. However, as we mentioned before, we can not tell causal relationships from the cointegration between variables; rather, we can only know the long-run correlations between them. Therefore, the negative correlation between overall crime rate and detection rate could be caused by two effects: 1) higher detection rate could reduce crime rate through deterrence and incapacitation effects; and 2) increased crime rate could disperse the limited law enforcement resource and thus reduce the probability of detection. Meanwhile, the crime-specific detection rates are positively cointegrated with the property crimes, namely burglary, theft and handling, and fraud and forgery. Such positive cointegrations are probably running through the crime rates to the detection rates: more crimes would require tougher crime combat policies, because it is unlikely that increased probability of detection would induce higher crime rates.

The detection rate is negatively cointegrated with the overall crime rate and positively cointegrated with individual property crimes as shown in table 3-8. Such long-term cointegrations can be explained as follows: increased property crimes require higher detections rates to combat and, therefore, part of police officers who used to work on other types of crimes could be re-allocated to solve property crimes. Such re-allocation of police personnel could lead to increases in the detection rates of property crimes and reductions in those of other crime types. As the overall detection rate is the proportion of overall crimes, including property and other types of crimes, which have been solved, higher detection rates for property crimes could lead to lower overall detection rate. Hence, while the crime-specific detection rates are positively cointegrated with individual property crime rates, the overall detection rate is, in the meantime, negatively correlated with the overall crime rate.

Secondly, the variable of people in custody shows negative cointegration with each type of crime rate and we argue that the causality could also be running either way. On the one hand, more people kept in custody could deter crimes through a signal of tougher punishment, and meanwhile, eliminate the possibility for prisoners to commit further crimes. On the other hand, higher crime rate could reduce the effectiveness and efficiency of the justice system. While the police could be less accurate in finding evidence when facing more reported cases, the court system could delay their sentences with more charged offenders. Therefore, the negative cointegration could also be caused by the negative effect of crime on custody rate.

Thirdly, the unemployment rate has positive integration with the overall crime rate, burglary as well as theft in equilibrium. This result implies that the motivation effect of unemployment has more significant impact on crime than the opportunity effect in long-run. Therefore, increased unemployment would lead to rising crime rates despite of it could also somehow reduce the crime opportunities at the same time. However, unemployment rate could be positively affected by crime rate through two channels: 1) in short-run, involving in crimes would reduce people's participation in legal sectors and thus lead to higher unemployment; and 2) in long-run, the criminal records of offenders would negatively affect their further payoffs from legal labour market as well as their probabilities of getting hired. Therefore, the positive cointegration between unemployment and crime (including overall, burglary, and theft) could be the result of two effects running through opposite directions. In contrary, unemployment has negative correlation with fraud and forgery, indicating increased unemployment is related to decreased fraud. This finding could be explained by the fact that a significant proportion of fraud and forgery is white-collar crimes and people need jobs to do so. An increase in unemployment implies fewer opportunities for white-collar crimes and lead to a reduction in fraud and forgery.

Table 3-7
Unit root tests for residuals

	ADF		PP	
Variable	Level	First Difference	Level	First Difference
Overall	-3.24***	-4.38***	-2.83***	-5.94***
Burglary	-4.97***	-5.78***	-4.69***	-7.50***
Theft and handling	-3.90***	-2.83***	-3.04***	-6.32***
Fraud and forgery	-4.06***	-4.73***	-3.09***	-6.08***

*** significant at 1% level; ** significant at 5% level; and * significant at 10% level.

Table 3-8
cointegration tests

	Overall crime	Burglary	Theft and handling	Fraud and forgery
Constant	3.62 (1.18)	0.89 (0.42)	1.79 (0.68)	-1.24 (1.12)
Detection rate	-0.16 (0.29)	0.11 (0.12)	0.18 (0.19)	0.38 (0.24)
Custody rate	-0.88 (0.27)	-1.36 (0.21)	-1.20 (0.19)	-1.13 (0.25)

	0.21	0.32	0.16	-0.01
ILO unemployment	(0.05)	(0.05)	(0.05)	(0.05)
	0.04	0.05	0.05	0.03
Linear trend	(0.01)	(0.01)	(0.01)	(0.01)

The values in the brackets are the estimated standard errors.

3.5.3 Dynamic Error Correction Models (ECM)

In the presence of cointegration, we are now able to estimate a dynamic ECM model using equation (3.4) for each crime. One of the conditions for a correctly specified ECM model is that we should include sufficient lags for the differenced variables so that the residuals would have the desired properties: normally distributed, not correlated, and having constant variances. We start our temptation by including two lags for the differenced variables to the right-hand side of equation (4) due to two reasons: 1) the sample size of 30 greatly restricts the potential to include large number of lags; and 2) as we are analysing annual data, including the information of two years before should be long enough to reflect the dynamics in normal cases. Our results suggest that, however, the ECM model is "over-fitted" while setting the lag length equal to two: although the obtained residuals have the right properties, the coefficients are broadly insignificant (including the ECM term). Thus, we reduce the lag length to one and re-estimate the system. This adjustment has made significant improvement: as some of the independent variables start showing significant coefficients (including the ECM term), the residuals still have the correct properties indicating the validation of the model. Therefore, based on the ECM model with the lag length of one, we try to derive a parsimonious dynamic model by dropping insignificant independent variables: one variable at a time until the adjusted R-squared starts falling and the standard error of regression starts rising.

As the results of parsimonious models are reported in table 3-9 below, we find that, firstly, the change in overall crime rate is most significantly affected by the contemporary change in detection rate: one percent increase in the probability of detection leads to a 0.89 percent reduction in the overall crime rate. Meanwhile, an increase in the custody rate also tends to reduce the overall crime, although this negative effect is statistically insignificant. Furthermore, the change in unemployment rate is positively correlated with the change in overall crime in short-run, reflecting the motivation effect of unemployment. However, the coefficient of differenced unemployment rate is also insignificant. As we incorporate the long-run information into the dynamic model, we find that the once-lagged error correction term has a negative and significant effect (at 5% level) on the growth rate of overall crime, consistent with expectation. This result indicates that, if the crime rate shifts away from its equilibrium level by one percent in current period, this deviation will be corrected by 0.43 percent in the next period. In other word, the speed of adjustment is around 40 percent per year.

The change in burglary is shown to be significantly affected by the change in custody rate as well as that of unemployment rate. Specifically, one percent increase in the number of people in custody will reduce burglary by 0.61 percent. Meanwhile, if the unemployment rate goes up by one percent, the burglary rate will accordingly rise by 0.47 percent. Consistent with expectation, the ECM term has obtained negative and significant coefficient at 1% level. It suggests a rather quick speed of adjustment: one percent drift from equilibrium in burglary will be pulled back to the level it should be by 0.94 percent in next period.

As seen in the section of data description, theft and handling exhibits identical trend over the examined period as burglary because they are both typical property crimes and share some common features. In this section, we find that their short-run changes also behave in similar way: the differenced theft and handling is significantly correlated with the changes of people in custody and unemployment. A one percent increase in the number of people in custody has a negative but bigger impact on theft than on burglary: 0.89 percent reduction will occur to theft accordingly. On the other hand, one percent increase in unemployment will lead to a 0.19 rise in theft, which is smaller than the response of burglary. The error correction mechanism shows that 63 percent of current disequilibrium will be corrected in next period. That is equivalent to say that one percent deviation from equilibrium in theft can be drawn back on track in less than two years.

As we finally move onto discussing fraud and forgery, we discover that the contemporary growth rate in fraud is positively correlated with its once-lagged value, and such correlation is significant at 1% level. This is implying that one percent growth in fraud one year before will result in a 0.54 percent growth in current period. Meanwhile, the coefficient of differenced custody rate tells that one point growth in the number of people in custody will reduce fraud and forgery by 0.56 point. In addition, the ECM term has a coefficient of -0.79, suggesting a rather strong adjusting effect: 79 percent of the disequilibrium will be diminished by the cointegration relationship.

We need to point out that we have applied a set of diagnostic tests on the residuals to examine their properties and the results are satisfactory. As shown in the lower part of table 3-9, we are able to accept that the estimated residuals are all normally distributed (only one exception being the residuals from the model of overall crime), serially independent and having constant variances. The only exception is that we reject the normal distribution for the residuals in the model of overall crime. In sum, our results of diagnostic tests suggest that the parsimonious dynamic models are valid and the estimated coefficients are not biased by either incorrectly behaved residuals or miss-specified equations.

Table 3-9

Parsimonious dynamic models

	D(overall)	D(burglary)	D(theft)	D(fraud)
Constant	0.03** (0.01)	0.03** (0.01)	0.03** (0.01)	0.02 (0.01)
D(crime)(t-1)	-	-	0.25 (0.17)	0.54*** (0.18)
D(detection)	-0.89*** (0.26)	-	-	0.29 (0.31)
D(custody)	-0.44 (0.27)	-0.61** (0.24)	-0.89*** (0.25)	-0.56* (0.29)
D(ilo)	0.13 (0.08)	0.47*** (0.07)	0.19** (0.08)	-
ECM(t-1)	-0.43** (0.18)	-0.94*** (0.15)	-0.63*** (0.17)	-0.79*** (0.20)
Adjusted R-squared	0.59	0.74	0.52	0.40
S.E. of Regression	0.05	0.05	0.05	0.05
Normality Test	8.50** [0.01]	0.20 [0.91]	2.64 [0.27]	0.72 [0.70]
Serial Correlation LM Test	0.16 [0.85]	0.02 [0.99]	0.03 [0.97]	0.72 [0.50]
White Heteroskedasticity	0.20 [0.99]	1.08 [0.42]	0.29 [0.98]	1.97 [0.14]
RESET Test	0.53 [0.48]	0.44 [0.51]	0.002 [0.96]	0.16 [0.69]

*** significant at 1% level; ** significant at 5% level; and * significant at 10% level. The values in the rounded brackets are the standard errors, while the values in the squared brackets are the associated *p*-values.

3.6 CONCLUSION

In this paper, we have tested the broadly concerned unemployment—crime relationship using annual time series data in England and Wales over the period 1971-2000. Accordingly, we have chosen the approach of cointegration analysis and error correction model to cope with the non-stationary variables. We have found that, in long-run, the overall and individual property crimes are cointegrated with unemployment as well as law enforcement instruments. Particularly, unemployment rate has positive cointegration with overall crime, burglary and theft, indicating that, for such crimes, the motivation effect is stronger than the opportunity effect. Increased unemployment rate would reduce the opportunity cost of committing crimes, as argued in Ehrlich (1973), and motivate potential offenders to engage in illegal activities. Such effect could offset the impact that higher unemployment would reduce the potential opportunities for property crimes. In contrary, unemployment is negatively correlated with fraud and forgery in equilibrium suggesting that higher unemployment could greatly affect the opportunities of while-collar crimes. Furthermore, while custody rate is negatively cointegrated with each crime rate confirming its negative effect as law enforcement instrument, detection rate has negative long-term correlation with overall crime and positive correlation with individual property crimes. We argue that the positive cointegration between detection rate and property crimes could be caused by a causality running from crime to detection: increased crime rate probably requires tougher crime control policies.

As we are also able to examine a dynamic error correction model, we find that the changes in crime rates are only affected by the contemporary changes in explanatory variables. Specifically, the change of custody rate has the strongest effect on the changes in crimes, which is constantly negative. The growth of unemployment also has relatively strong and constant effect: it is positively correlated with overall crime, burglary, and theft and handling. This finding is consistent with their long-run relationships.

The error correction term is a beneficial feature for using cointegration and ECM analysis. Its associated coefficients have shown rather quick adjusting process. The highest speed of adjustment occurs to burglary indication 94 percent of disequilibrium will be corrected in one year's time. Meanwhile, the overall crime has the lowest speed of adjustment and 43 percent deviation from equilibrium will be pulled back in the next year.

We have adopted Engle-Granger two-step procedure to detect for cointegration. This approach has a few limitations. Firstly, cointegraion tested by this approach is unable to tell which variable should be taken as explained and which ones are explanatory. Theoretically, with infinite samples, treating different variables as dependent should yield the same cointegration relationship. However, with limited sample size, using different dependent variables is possible to generate different correlations. Secondly, this technique ignores the possibility of more than one cointegration relationship in the case of including three or more variables. And thirdly, as this approach takes two steps to perform, a mistake involved in the first step will be automatically carried into the second stage and thus generate misleading results. Such potential issues associated with Engle-Granger two step procedure can be avoid by using Johansen cointegration test. However, Johansen technique could induce the identification problem. When there are more than one cointegration relationships, it would be rather difficult, if not impossible, to identify which is the "true" long-term relationship.

Our cointegration model includes four variables as well as a constant and a linear time trend, as defined in equation (3.1). Therefore, we cannot ignore the possibility of multiple cointegration vectors. Consequently, we have applied both eigenvalue and trace tests to detect the number of cointegrating vectors for each crime type. Our results of tests suggest that there only exists one cointegrating relation among relevant variables for the overall crime, burglary and theft and handling, according to both eigenvalue and trace tests. On the other hand, trace test indicates one cointegration relationship for fraud and forgery while the eigenvalue test implies zero. Since the cointegration tests have eliminated the possibility of more than one cointegration vectors for each type of crime, our application of Engle-Granger two step approach will not be affected by the existence of multiple cointegrating relationships.

An issue that usually concerns with time series analysis (and panel data analysis in some cases) is the potential structural break, particularly when the data covers several decades. Our cointegration analyses in this chapter are conducted upon the time series data over 30 years (1971-2000). During this period, any significant social or economic change could potentially affect the correlations between our concerned variables. For instance, one of the most well-known events in the examined period is that Margaret Thatcher became the leader of the Conservatives in 1979 and implemented her policies of reducing state intervention and encouraging free market. Such polices dramatically increased the unemployment rate in the UK by privatizing many nationally-owned enterprises. Similarly, the economic recession in the early 90's also caused awful labour market conditions. There are methods to control the influences of such significant events and one of them is to carry out structural break test. Alternatively, we could reduce the probability of structural break by shortening the time period being covered. As shown in our later chapters, we examine much shorter time periods in our panel data and spatial analyses.

Another potential shortcoming of this chapter is that we have only 30 samples in total, which has greatly restricted our selection of explanatory variables. However, limited sample size is a common problem for using time series data. In order to reduce the small sample bias, we have tried to strictly control the number of explanatory variables and only included the most important ones. Aside from unemployment which is our main concern, we argue that law enforcement variables should have the closest correlation with crimes. Our argument is supported by most empirical papers because, as long as they include variables such as detection rate, conviction rate, imprisonment rate etc., these variables show negative and significant effects on crimes most of the time. Therefore, we have included crime-specific detection rate and custody rate along with unemployment in order to control for their correlations with crimes. In the next chapter, we will tackle the issue of limited sample size by employing panel data analyses.

Chapter Four: Panel Data Analyses

4.1 INTRODUCTION

The economic analysis of crime is normally traced to Becker (1968) whose seminal work on crime and punishment has been the starting point of almost all works in this area. Ehrlich (1973) extends Becker's model and looks at economic factors accounting for changes in the incentives for crime. These early analyses provide a framework for identifying factors that affect crime. Much empirical work in this area uses variables, such as unemployment, income level, that are believed to affect the costs and benefits of crime. These variables are presumed to be the factors that affect rational criminal behaviour. The main aim of these papers (including this work) is to verify if these "deterrence" variables do affect criminal behaviour.

In this chapter, we examine the factors that we believe contribute to crime in England and Wales using panel data for the period 1992-2005. Various factors have been identified in the literature as affecting crime rates. We do a rigorous empirical analysis of these factors for various types of crime in England and Wales for the period under consideration. By doing this, we seek to identify what types of crime respond most to the so called economic factors and what do not. This is clearly useful for policy purpose.

While there are only a limited number of literature in England and Wales analysing factors affecting crime rates, even fewer studies have done with panel data analysis at our level of disaggregation. In our review of the literature in the next section, we find some common issues that these papers do not handle and we intend to fill the gap in this work. First of all, most of the literature we have reviewed focus on analysing property crimes such as burglary and theft (e.g. Witt *et al.*, 1998; Reilly and Witt, 1996). This is probably because people usually believe that property crimes are more responsive to the factors affecting the costs and benefits of committing crimes, such as law enforcement variables and labour market opportunities. However, we argue that violent crimes, such as sexual offences and personal violence, could also be analysed in the same framework. Although violent crimes are normally not committed for economic benefits, they should still be influenced by variables that change the "price" and opportunities for doing so.

Our second finding on the existing literature in England and Wales is that most of them define their aim as identifying the effect of unemployment on crime. They, nevertheless, have failed to include a complete set of crime-influencing variables in their empirical modes. Both Becker (1968) and particularly Ehrlich (1973) have theoretically demonstrated the effects of various factors on crime such as law enforcement, social-economic status, demographic composition and so on. Omitting some of these variables could not only lead to mis-specified model, but also bias the results if the omitted variables are correlated with the included explanatory variables.

The third shortcoming of some existing literature is that they have ignored the potential endogeneity of law enforcement variables. This issue is essential in detecting the relationship between crime and its relevant factors, as without a control for endogeneity, the reversed causality from crime to crime control variables could generate biased estimation and thus misleading implications.

In this chapter, we intend to empirically test the correlation between different types of crime rates and their influencing factors while controlling for the previously mentioned issues. We break down the total recorded crime rate into eight individual categories as defined by the Home Office and analyse six of them with our empirical framework. They are violence against the person, sexual offences, robbery, burglary, theft and handling, and fraud and forgery. While the first three types are defined as violent crimes, the last three categories are property crimes. We try to avoid the omitted variable bias by choosing our explanatory variables according to Ehrlich (1973). In particular, we specify an empirical model where crime rate is affected by law enforcement instruments, social-economic status as well as demographic composition. Moreover, our use of panel data allows us to control for unobserved heterogeneity which can further reduce the risk of mis-specified equation. We control for the endogenous law enforcement variables by adopting the GMM technique and applying instrument variables. Further, we also allow the crime rate to depend on past crime rates as we believe that crime rate could have shown persistence over time due to either the recidivism of offenders or the effects of lagged explanatory variables, or both.

Our panel data is disaggregated by police force areas in England and Wales over the period 1992-2005. As a unique feature, we also apply a different, but overlapping data set, covering the time period 1987-2005 in England and Wales to check the robustness of our results. The difference between the two data sets, aside from the period being covered, is that the independent variables of unemployment rate and real average weekly earnings are disaggregated by police force areas in the data set of 1992-2005 while they are disaggregated by regions in the data set of 1987-2005. As the shorter data (1992-2005) are more accurate in reflecting the variations in explanatory variables, we report the results generated by this data set as our main finding. We use the obtained results generated by analysing the longer data date through the same procedure as our robustness check.

Our main findings include, first of all, the property crimes are better explained than the violent crimes by our empirical model which is derived according to Ehrlich (1973). Different crimes do react differently to the changes of incorporated explanatory variables. Secondly, among all the explanatory variables, law enforcement variables have the strongest impacts on different types of crime. Their negative and significant associations with crimes have confirmed their deterrence and incapacitation effects as predicted by theoretical models and such results are the most important findings of the chapter. Thirdly, the social-economic factors, such as unemployment and real earnings, have mainly picked up their opportunity effects on crimes, particularly property crimes, suggesting worse social-economic conditions will reduce crime due to few crime opportunities. Furthermore, our results are rather robust when applying the data set 1987-2005 with only few exceptions, which we will provide an explanation for.

We structure this chapter in the following order. Section two reviews empirical literature in England and Wales as well as other countries that investigate similar questions using panel data. In section three, we firstly discuss the theoretical background based on which we construct our empirical model. Next, we specify the empirical model we will test and introduce the econometric methodologies which we shall adopt. Whilst section four describes the two data sets in details and presents summary statistics, section five reports and examines our estimation results. As there are two sets of data being employed, the results generated by different data sets will be compared together to show whether they are robust across data sets. Finally, section six summarises the main findings of this work and points out the potential shortcomings that could be improved in the future.

4.2 LITERATURE REVIEW

We construct our empirical model based on the theoretical frameworks developed in Becker (1968) and Ehrlich (1973). As these works have already been discussed extensively in the chapter of literature review, we will only briefly outline both of them in the section on theoretical background and estimation methodology after which we introduce our empirical model.

In this section, we will firstly review the literature in England and Wales that investigate the determinants of crime using panel data. These papers are important because they are most closely related to our work. Next, we extend our discussion to similar literature from other parts of the world, such as Scotland, France, Greece as well as the U.S., in order to briefly introduce their research on the same topic by implementing panel data. Finally, we point out the potential weakness of the literature, particularly in England and Wales, which we try to overcome in our work.

4.2.1 *Empirical Applications*

4.2.1.1 U.K. Studies

Carmichael and Ward (2000; 2001) are interested to test the relationships between youth unemployment and different types of crime rates in England and Wale using panel data. The main difference between the two works is that Carmichael and Ward (2000) disaggregate their panel data on regional level over the period 1985-1995, while their later work apply an extended data set disaggregated on county-level covering the years 1989-1996. The earlier work, Carmichael and Ward (2000), uses the crime rates of burglary, criminal damage, robbery, theft, violence against the person, and total crime as dependent variables. As they are primarily interested in the unemployment – crime relationship, they separately estimate the effects of youth and adult unemployment rates on crime to investigate whether crime rates would respond differently to them. They also control for various independent variables including the percentage of white population, crime-specific clear-up rate, the percentage of convicted criminals receiving prison sentence, and the average sentence length. In order to cope with the panel data structure, they adopt the OLS regression with region-specific fixed effects to take into account the regional unobservable heterogeneity. Their results suggest that, while there is positive correlation between unemployment and crime in general, different crime rates could be affected differently by youth and adult unemployment. Specifically, burglary is the only crime positively affected by both youth and adult unemployment rate. Criminal damage and robbery only have positive correlations with youth unemployment rate, while theft is only positively related to adult unemployment rate. In contrary, violence against the person shows no significant correlation with either youth or adult unemployment rate.

In the latter work (2001), as well as renewing their panel data by upgrading the level of disaggregation, Carmichael and Ward have also made some adjustment to the model specification. On the one hand, they include the crime rate of fraud and forgery in their analytical framework as dependent variable. On the other hand, they expand their set of independent variables by incorporating population density and the percentage of births outside marriage to capture the degree of urbanization and the traditional family values respectively. Based on the same estimation methodology, both youth and adult unemployment rates have shown positive and significant correlations with burglary, theft, fraud and forgery, and total crime. Only adult unemployment rate is significantly related to robbery, and neither youth nor adult unemployment has shown significant effect on criminal damage and violent crime.

Witt *et al.* (1998; 1999) also adopt similar research strategy: while they use regional level panel data over the period 1979-1993 in the earlier paper, they further disaggregate the data by police force areas for the years 1986-1996. In both articles, they seek test the relationship between unemployment and different crime rates in England and Wales. Moreover, they also intend to identify the effect of income inequality on crime rates. In the earlier work, the crime rates being analysed include burglary, theft from a vehicle, other theft, shoplifting, and robbery. They assume these crimes are affected by the explanatory variables of wage inequality, unemployment rate, population density, police employees, as well as demographic variables controlling for age distribution. In order to eliminate region-specific fixed effects, their analyses are based on first-differenced OLS estimations. The results show that both variables of their interest have significant correlations with the property crimes. In particular, the growth of male unemployment has positive and significant effect on all the five crime rates being studied. Furthermore, the wage inequality is also positively correlated with all the five crime rates, and among which burglary, theft from a vehicle and robbery are the most responsive types to the increase in wage inequality.

Witt *et al.* (1999) have made significant changes in their model specification from their earlier paper. While the dependent variables are burglary, vehicle crime, handling stolen goods as well as other theft, the independent variables include the once-lagged crime rate being analysed, unemployment rate, wage inequality, cars per capita, once-lagged police per capita, as well as year-specific and area-specific dummies. As presented in the paper, they allow the crime rate to follow an AR(1) process because the lagged crime rate could reflect the tendency of criminals to keep committing crimes even after the other crime-influential factors have changed. Furthermore, they adopt the generalised method of moments (GMM) technique to cope with the inclusion of lagged dependent variable. Their results show that, firstly, the unemployment rate has positive and significant coefficient in the analysis of each type of crime, which is consistent with the finding in Witt *et al.* (1998). Also, the wage inequality has constantly shown positive and significant effect on crime rate in all cases. This result differs from Witt *et al.* (1998) in which the wage inequality only has significant effect on certain types of crime.

Reilly and Witt (1996) aim to identify the effects of unemployment as well as law enforcement instruments on property crimes in England and Wales. The panel data employed in their analysis is disaggregated by police force areas covering the period 1980-1991. The dependent variables include burglary, theft and robbery, and each of them is predicted by the explanatory variables of crime-specific clear-up rate, average sentence length, and male unemployment rate. The basic analysis is conducted by applying the OLS estimation incorporating both area-specific and year-specific fixed effects. However, this basic analysis is unable to control the simultaneity between crime rates and law enforcement variables. Accordingly, they apply an unrestricted error-correction mechanism (ECM) model to solve for the endogeneity of clear-up rate and average sentence length. Their investigations mainly reveal that, in long-run, the average sentence length has negative effect on burglary and robbery while the clear-up rate is negatively affecting burglary and theft. The negative effects of both average sentence length and clear-up rate are consistent with the predicted deterrent effect as law enforcement variables. On the other hand, in short-run, the average sentence length only negatively affects theft while the clear-up rate only negatively affects robbery. As one of their main concerns, unemployment rate has exhibited positive effect on both burglary and theft suggesting that the motivation effect of unemployment is stronger than its opportunity effect.

Machin and Meghir (2000) exclusively investigate the relationship between property crime rates and worsened labour market conditions of less skilled workers. To do so, they utilise panel data by police force areas in England and Wales over the period 1975-1996. The regression analysis is initially applied on the aggregated property crime as well as vehicle crime. Then the aggregated property crime is broken down into burglary and theft and handling. Each type of crime is explained by the 25th percentile real hourly wage, the percentage of people aged 15-24 as well as the conviction rate. In addition, both area-specific and year-specific dummies are included all the time. The undertaken estimations have found that, firstly, the 25th

percentile real wage has negative correlation with the aggregated property crime rate. After breaking down the aggregated property crime into sub-categories, this variable is still negatively correlated with vehicle crime, burglary, and theft and handling. As the 25th percentile real wage is assumed to measure the wage rate of low skilled workers, the previous results advocate that higher incomes for low skilled workers will reduce the occurring of property crimes. Moreover, the conviction rate constantly displays negative correlation with both aggregated and broken-down property crimes. This result is supportive for the expectation that conviction rate, as proxy for the probability of punishment, has a deterrent effect on property crimes.

4.2.1.2 European Studies

Apart from the literature in England and Wales that investigate the relationships between crime rates and relevant factors, similar question has been addressed for other European countries. By way of example, Reilly and Witt (1992) try to examine the relationship between unemployment and crime in Scotland using panel data on regional level over the period 1974-1988. The dependent variable is the overall crime rate measured by the number of offences per 100 population. The explanatory variables include the unemployment rate and the number of completed public authority houses per capita. While the former indicates the labour market opportunity, the latter is used to capture the regional influence of government or local authority due to the unavailability of regional level government expenditure figures. In addition to the standard OLS estimation, they also apply fixed and random effects models in order to get rid of the region-specific factors. Furthermore, Cochrane-Orcutt and Prais-Winsten procedures have also been implemented to deal with the potential cross-sectional correlations in the error terms. The most noteworthy finding of this paper is the robust positive correlation between unemployment rate and crime rate across different estimations. The coefficient of unemployment is significant in all cases while the magnitude ranges from 0.15 to 0.35. Whereas, the effect of public houses per capita on crime is sensitive to estimation method. The results show that, once the region-specific effects are incorporated through either fixed or random effects model, the public houses per capita exhibits negative and significant correlation with crime which is consistent with expectation.

Pyle and Deadman (1994b) re-estimates the model specified in Reilly and Witt (1992) by extending the panel data to the period 1974-1991. By adopting the same explanatory variables and estimation methods, Pyle and Deadman argue that the previously positive correlation between unemployment and crime becomes insignificant in both fixed and random effects models. Besides, the coefficient of unemployment has substantially smaller magnitude in the analyses using the extended data. The reason behind this reduced correlation between unemployment and crime may be that, as pointed out by Pyle and Deadman, both unemployment and crime have common upward time trends over the period 1974-1988 which is reflected by the previous positive correlation between the two variables. However, simply extending the data set by three years have changed the common trend over time between unemployment and crime and weakened their estimated correlation.

Another example of empirical study on a European country is Edmark (2003), which uses panel data on county-level from Sweden over the years 1988-1999 to explore the effect of unemployment on property crime rates. The property crime rates under scrutiny are burglary, robbery, car theft, bike theft, theft/pilfering from motor vehicle and shop, and fraud. These crime rates are separately explained by a set of independent variables including unemployment rate, clear-up rate, average income, the proportion of divorced, population density, the proportion of people with higher education, the proportion of people on social allowance, the proportion of foreign citizens, the proportion of 15-24 years old, and the sales of alcohol at the National Liquor Monopoly. In addition to these controlling variables, both county-specific and year-specific fixed effects are included in the estimations to count for the unobservable features that would be otherwise correlated with the independent variables. Furthermore, he also makes an attempt to control for unobservable county-specific time trends by adding both linear and quadratic time trends into the empirical model. The estimated results have shown strong evidence that unemployment is correlated with some property crimes. Particularly, unemployment rate is positively correlated with burglary and car theft and such correlations are insensitive to different model specifications. In the meantime, unemployment rate has shown positive effect on bike theft in the model without time trends as well as the model with both linear and quadratic time trends. Furthermore, the clear-up rate has been a strong predictor for the property crime rates. It has constantly shown a negative effect on each type of crime rate under study. However, as Edmark acknowledges, one of the potential issues associated with this model is that the model is unable to control for the simultaneous relationship between the clear-up rate and crime.

A most up to date examination on the relations between crime rates and potential explanatory factors is Saridakis and Spengler (2009). The article employs regional level panel data in Greece over the period 1991-1998 to examine the issue of their concern: the relationship between crime, deterrence and unemployment. The focus their analyses on the crime rates of breaking and entering, theft of motor cars, robbery, murder, serious assault, and rape; and each of the crime is regressed on the once-lagged crime rate, clear-up rate and unemployment rate. Instead of including region-specific dummies, they take first differences for both dependent and independent variables to eliminate the region-specific fixed effects. In order to correct the estimation biases caused by including the once-lagged dependent variable, the GMM technique is executed to employ instruments for the lagged crime rate. Their results denote that property and violent crimes respond quite differently to the controlling variables. While the clear-up rate has negative effect on property crimes including robbery, breaking and entering, and theft of motor cars, it is not significantly correlated with any of the violent crimes. Similarly, while the unemployment rate has positive correlation with all the property crime rates, none of the violent crimes are significantly affected by unemployment.

Fougere *et al.* (2003) are interested in examining the influence of unemployment on crime rates in France using panel data on regional level over the period 1990-2000. Their analyses are performed by applying the OLS estimations including both region-specific and year-specific fixed effects. The dependent variables under investigation are 17 types of crime rates, including both property and violent crimes.[21] Each type of crime rate is regressed on unemployment rate, which is the main concern of their paper, as well as social-demographic controls.[22] In general, results obtained suggest that property crimes are better explained than violent and family crimes by the same

[21] The crime rates being analysed are armed or violent robberies, burglaries, car thefts, motorbike thefts, thefts of objects from cars, shoplifting, pick pocketing, receiving stolen goods, homicides, voluntary wounds, blackmails and threats, rape and other sex offences, family offences, drug offences, damage to vehicles, illegal weapon ownership, and violence against police.

[22] The social-demographic variables include fraction of foreigners from North Africa, fractions of people aged 15-24 and 25-49, fraction of men living alone, fraction of people in single-parent families, fraction of people without any diploma, fraction of high school graduates, fraction of people living in rural areas, fraction of people living in cities between 20,000 and 200,000 inhabitants, fraction of people living in cities with more than 200,000 inhabitants, and fraction of people living in Paris and its suburbs.

explanatory variables. Being the main interest of this paper, unemployment has shown negative correlation with property crimes including burglaries, theft crimes, as well as drug offences. In contrary, unemployment is detected to be positively correlated with violent crimes of homicides, threats, violence against police. After unemployment is broken down into age groups, youth unemployment has shown positive impact on most crimes whereas unemployment of older age groups has negative impacts on most crimes.

4.2.1.3 U.S. Studies

Meanwhile, Doyle, *et al.* (1999) and Gould, *et al.* (2002) defined their interests as detecting the effects of labour market conditions on crime in the United States and both papers have been previously offered an extensive discussion in Chapter Two. Doyle, *et al.* (1999) use state-level panel data over the years 1984-1993 to examine aggregated property and violent crimes. They assume that either property or violent crime is predicted by explanatory variables including real annual wage, unemployment rate, Gini coefficient, percentage of young male aged between 15 and 29, probability of arrest and per capita police officers. Moreover, the panel data structure also enables them to include state-specific effects to count for the unobservable factors that differ across states. In order to control for the endogeneity of arrest rate and per capita police, they adopt the GMM technique and apply instruments for both endogenous variables. Their principal finding is that, whilst real annual wage is negatively correlated with both property and violent crimes, Gini coefficient is largely insignificant in explaining both of them.

Gould, *et al.* (2002) has further refined their focus on the correlation between crime and the labour market conditions of less-skilled men, which, in this article, is defined as non-college-educated male. Based on county-level panel data over the years 1979-1997, they regressed each index crime rate on the independent variables including weekly wages, unemployment rate, per capita income, arrest rate, state expenditure per capita on police, and state police employment per capita. Additionally, they also control for the changes in the age, sex and race composition as well as county-specific and year-specific fixed effects. They have largely found that, whilst the wage rate of the less-educated is negatively correlated with both property and violent crimes, the unemployment rate of less-skilled men have positive effect only on property crimes.

4.2.2 Weakness of Existing Economic Literature

Most studies introduced above have shown their interests in examining the relationships between unemployment and different crime rates, and most of them have indeed found certain evidence supporting a positive effect of unemployment on crime. Nevertheless, some common unsolved issues in some of the existing literature need to be outlined so that it can be rectified in our analyses in this chapter.

Firstly, by taking the unemployment – crime relationship as their primary concern, several current studies have failed to control for a relatively complete set of crime-influential variables in their analyses. As suggested in Becker (1968) and Ehrlich (1973), crime rates should be affected by factors of different aspects such as law enforcement, social-economic conditions, as well as demographic compositions. Some literature, however, have only included part of the relevant factors and omitted other potentially important ones. For example, Carmichael and Ward (2000; 2001) have omitted the legal income from labour market which may reflect the attractiveness of legal activities; Witt *et al.* (1996; 1999) do not incorporate the legal income as well as the expected severity of punishment in both papers; Reilly and Witt

(1996) has left out both legal income and demographic factors whilst Reilly and Witt (1992) do not control for legal income, demographic compositions, as well as crime control instruments. Therefore, as a result of omitting important factors from the empirical model, the estimated correlations between crime rates and the variables of concern, mainly unemployment rate, could be spurious. The positive correlation between unemployment and crime, as proposed in some articles, may be due to the positive correlation between crime and a third variable which is neglected from the specification but correlated with unemployment rate.

In chapter, we attempt to introduce a relative complete set of explanatory variables, adherent to Becker (1968) and Ehrlich (1973). We take into account the law enforcement instruments by using detection rate and the prison population as proxies for the probability of apprehension and the severity of punishment. At the same time, we measure the labour market opportunities by income inequality, unemployment rate and real average weekly earnings. In addition, the proportion of people aged 15-24 is also incorporated as the demographic control to capture the share of more crime-prone population. Furthermore, we allow the crime rates to follow an AR(1) data generating process by incorporating once-lagged crime rates as explanatory variables. We assume the lagged crime rates to capture the persistence in crime as well as the effects of lagged explanatory variables. By employing panel data, we can further reduce the probability of omitted variable bias through eliminating the unobserved area-specific fixed effects.

Another problem appears in current literature is that the law enforcement variables are not treated as endogenous (e.g. Carmichael and Ward, 2000; 2001; Witt *et. al,* 1998;

Edmark, 2003). As neglecting the endogeneity of law enforcement controls will break down the consistence of estimation, we will explicitly treat our law enforcement variables as endogenous by implementing instrumental variables. Such strategy is achieved by carrying out the GMM estimations which will be introduced in latter part.

4.3 Theoretical Background and Estimation Methodology

We construct our empirical model heavily drawing from the theoretical frameworks developed in Becker (1968) and Ehrlich (1973). Frist, we will sketch how crime is affected by different factors according to these two articles. Next, by assuming a linear function, we spell out our empirical model accordingly and discuss the associated estimations issues.

4.3.1 Theoretical Discussion

Becker (1968) relates the number of offences one would commit with his probability of apprehension and the severity of punishment by assuming each individual is economically rational and trying to maximize his expected utility from committing crimes. The expected utility from committing an offence for individual *i* is defined as

$$EU_i = p_i U_i(Y_i - f_i) + (1 - p_i) U_i(Y_i)\,, \tag{4.1}$$

where Y_i is his income from an offence; U_i is his utility function; p_i and f_i respectively represent his probability of apprehension and the monetary equivalent of punishment. Therefore, it can be shown that one's expected utility from an offence is reduced by either higher probability of apprehension or more severe punishment, as demonstrated by the first-order conditions of the expected utility with respect to the probability of apprehension and the severity of punishment respectively

$$\frac{\partial EU_i}{\partial p_i} = U_i(Y_i - f_i) - U_i(Y_i) < 0$$

and

$$\frac{\partial EU_i}{\partial f_i} = -p_i U_i^{'}(Y_i - f_i) < 0 \, .$$

Thus, the number of offences one would commit can be related to his probability of apprehension and the severity of punishment as given by equation (4.2).

$$O_i = O_i(p_i, f_i, u_i) \tag{4.2}$$

Furthermore, O_i is expected to have the following properties:

$$O_{p_i} = \frac{\partial O_i}{\partial p_i} < 0$$

and

$$O_{f_i} = \frac{\partial O_i}{\partial f_i} < 0 \, ,$$

which imply that the number of offences one would commit is decreasing as either the probability of apprehension or the severity of punishment increases.

As an extension to Becker (1968), the model in Ehrlich (1973) is also developed upon the assumption that individuals are utility maximizing. By allowing individuals to freely allocate their time between legal and illegal activities, people’s labour market conditions have been incorporated in the decision of whether or not to commit crimes. Specifically, both the probability of apprehension and the severity of punishment are measuring the risk of illegal activities and an increase in either of them would reduce one’s expected return from illegal activities. On the other hand, the unemployment rate and legal income level are respectively measuring the uncertainty and potential return of legal activities. Each individual is assumed to maximize his expected utility by optimally allocate his time and other resources between legal and illegal activities. Thus, one's expected utility function to be maximized can be specified as

$$\begin{aligned} EU(X_s, t_c) &= (1 - p_i)(1 - u_l)U(X_a, t_c) + (1 - p_i)u_l U(X_b, t_c) \\ &+ p_i(1 - u_l)U(X_c, t_c) + p_i u_l U(X_d, t_c) \end{aligned}, \qquad (4.3)$$

where U is his utility function; t_c is his leisure time; p_i is his probability of apprehension; u_l is the unemployment rate; and Xs are his monetary returns from four status that he could be end up with[23].

By maximizing the utility function, one's participation in illegal activities can be related to the included factors. First, an increase in either the probability of apprehension or the severity of punishment will reduce one's incentive to participate in illegal activities because the expected cost of punishment of doing so becomes higher. Second, either increased illegal payoffs or decreased legal incomes will increase one's participation in illegal activities due to the increased relative benefit between illegal and legal activities. Third, an increase in unemployment rate will have ambiguous effect on one's participation in illegal activities. On the one hand, higher unemployment rate will unambiguously increase one's participation in illegal activities through reducing his opportunity cost of doing so. On the other hand, an increase in the probability of the least desirable status (unemployed in legal activities and failed in illegal activities) will increase one's the demand for wealth and hence reduce his incentive to participate in illegal activities.

[23] The four statuses one could end up with are *a*. successful in illegal activities and employed in legal labour market, *b*. successful in illegal activities and unemployed in legal labour market, *c*. failed in illegal activities and employed in legal labour market, and *d*. failed in illegal activities and unemployed in legal labour market.

Given the crime participation on individual level, the aggregated supply of offences can be specified as a function of the factors influencing individual crime participation

$$Q = \Psi(P, F, W_i, W_l, U, \Pi), \tag{4.4}$$

where Q is the aggregated supply of offenses; P and F respectively represent the average values of probability of apprehension and severity of punishment across all individuals; W_i and W_l respectively represent the average illegal and legal returns across all individuals; P and Π respectively represent the average unemployment probability and environmental factors, such as family background, education and so on, across all individuals. We expect the aggregated level of offences is affected the same way by the relevant factors as the individual crime participation.

4.3.2 Empirical Model and Estimation Methodology

Based on equation (4.4) above, we specify our empirical model to be tested in the form of equation (4.5) below by assuming a linear relationship between crime rate and its influencing variables.

$$\begin{aligned}
&\ln(crime)_{it} = \alpha_i + \beta_0 \ln(crime)_{i(t-1)} + \beta_1 \ln(\det ection)_{it} + \beta_2 \ln(prison)_{it} \\
&+ \beta_3 \ln(Gini)_{it} + \beta_4 \ln(young)_{it} + \beta_5 \ln(unemployment)_{it} + \beta_6 \ln(earnings)_{it} \,. \quad (4.5) \\
&+ dummy + trend + \varepsilon_{it}
\end{aligned}$$

In our specification, each type of crime rate is assumed to be affected by law enforcement instruments, labour market opportunities, demographic composition, as well as its once-lagged value.

We use the detection rate and prison population as proxies for the probability of apprehension and the severity of punishment and expect both of them to have negative effects on crime rate as predicted by the theoretical models. As we have realised that some papers use conviction rate as a proxy for the probability of apprehension, we argue that the deterrence effect of detection rate on potential offenders should be at least as strong as that of conviction rate. This is because, for real offenders, being detected is often regarded as the first step of punishment. Even though the conviction depends on a number of exogenous variables such as the evidence presented by the police, who the judges are etc., the potential criminals still would try to avoid detection in first place. Very few offenders will be so confident, when arrested, that they will not be convicted by the court. We use prison population to measure the severity of punishment because, firstly, the other commonly used variable, the average sentence length, is not separately provided for individual crime types; and secondly, prison population can be used as an alternative as its deterrence effect on crime has been demonstrated by a number of studies such as Levitt (1996), Wolpin (1978), and Saridakis (2008). Prison population could be an imperfect measure for the severity of punishment. However, other things being equal, more people in prison in time period *t* comparing to previous period could mean that offenders are getting released less often or, for the same type of crime, more people are sentenced into prison. Therefore, on average, the severity of punishment increases.

The labour market conditions are represented by Gini coefficient, unemployment rate and real average weekly earnings in our empirical model. Gini coefficient, measuring income inequality, is expected to be positively correlated with crime rates (Choe 2008; Kelly 2000; Scorzafave and Soares 2009). This is because higher Gini coefficient indicates that larger proportion of national income is possessed by a smaller group of people and the majority of the population disproportionally shares the rest of the national income. In such case, more people are at the bottom end of wealth distribution with lower opportunity cost if commit crimes. Consequently, this may lead to higher crime rates. The effect of unemployment on crime is ambiguous to predict as argued in Ehrlich (1973). Furthermore, unemployment rate may have both motivation and opportunity effects on crime as argued in Cantor and Land (1985). As higher unemployment motivates potential offenders to commit crime by decreasing their opportunity cost of doing so, it also reduces the opportunities for certain crimes and thus tends to reduce those crime rates. Therefore, the net effect of unemployment rate will depend on which effect is stronger, motivation or opportunity, and could be positive, negative or even insignificant. Likewise, the average weekly earnings is also expected to have ambiguous effect on crime due to the same reason. Whilst higher earnings could reduce people's incentives to commit crimes, it could also increase the opportunities for property crimes. Therefore, the net effect of real earnings on crime rates could be positive, negative, or insignificant.

Our empirical model includes the percentage of population aged 15-24 for demographic control because young people are usually regarded as more crime-prone than older age groups due to their lower opportunity cost (Levitt 1988; 1998 and 1999; Cohen and Land 1987). Furthermore, the punishment for young offenders under the age of 18 is much lenient than that for adult offenders. Moreover, the criminal records of juvenile offenders will be sealed when they reach the age of 18 thus not affecting their future labour market outcomes. For these reasons, we expect a positive correlation between the proportion of young people and the crime rates. However, we could also argue that young people have higher opportunity cost of committing crimes in terms of future job opportunities than older age groups. Particularly for young people over the age of 18, committing crimes and getting caught could reduce their future opportunities in labour market because they are less likely to get hired and could be paid with lower wages on average. Such effect could deter young people from participating in illegal activities. Therefore, in theory, if this effect dominates, the proportion of young people could have negative effect on crimes.

We include the once-lagged crime rate as explanatory variable to measure the persistence of crime rate over time, and there could be several reasons why crime rate can be thought to be self-correlated over time: (1) recidivism caused by, among other things, negative expected payoffs from the labour market for being a criminal; and (2) business cycle features such as recessions affecting the crime rate over successive periods. Furthermore, the lagged crime rate could act as a proxy for the effects of lagged independent variables, such as lagged unemployment rate, lagged detection rate, and lagged earnings and so on, that may explain the contemporary crime rate. Therefore, we expect the coefficient of once-lagged crime rate to be positive.

As shown in equation (4.5), we count for area-specific unobserved factors by incorporating the areas-specific fixed effects. However, given that prison population and Gini coefficient are only available on national level, we are unable to apply year-specific dummies to eliminate the factors that only vary by years. For compensation, we adopt time trend instead to count for unobserved trends. The dummy variable included in the empirical model indicates the counting rules change introduced in 1998, which will be given detailed discussion in the section of data description.

We firstly estimate the empirical model using the standard OLS method without controlling for lagged crime rate and area-specific fixed effects as an exploratory analysis. The potential problems associated with this approach are that, firstly, the OLS estimation could be biased by the potential correlations between the independent variables and time-invariant police force area characteristics $\in_i$. Such area-specific characteristics, if correlated with the independent variables and not eliminated from estimation, will be left in the error term and cause correlation between the error term and the independent variables. Thus, the OLS estimation will be biased by violating the assumption that explanatory variables should be independent from the error term.

In order to avoid the potential correlation between the independent variables and the error term, we re-estimate our empirical model with cross-sectional fixed effects method that explicitly eliminates the area-specific characteristics. We adopt this approach in order to show how the coefficients are affected by eliminating the area-specific fixed effects. However, in this stage, we still exclude the once-lagged crime rate from the equation because its inclusion requires special treatment.

The second problem associated with the OLS estimation, as well as the fixed effects model, is the potential reverse causality from crime rate to certain explanatory variables. In particular, a positive shock in the crime rate may, in short-run, reduce the detection rate and prison population if the budget for law enforcement is inflexible. In long-run, on the other hand, increased crime rate could trigger higher investment by the government on police and justice expenditure which, in turn, may increase detection rate and prison population. This would again present a rise to the correlation between the error term and the law enforcement instruments, detection rate and prison population, and thereby bias both the OLS and cross-sectional fixed effects estimations.

Furthermore, our inclusion of the once-lagged crime rate as explanatory variable will cause the correlation between itself and the area-specific fixed effects $\in_i$. A common strategy to eliminate the correlations between the fixed effects and independent variables is to estimate a fixed-effects or first-differenced model. Conversely, the inclusion of the lagged crime rate $crime_{i(t-1)}$ as an explanatory variable invalidates the conclusions from such traditional panel data estimation techniques (Baltagi 2004; Hsiao 2003). Instead, we estimate our full empirical model given by equation (4.5) using a dynamic panel data fixed effects estimation strategy developed in Arellano and Bond (1991). The main intuition behind such estimation methodology is briefly outlined below.[24]

For the ease of exposition, the empirical model can be re-written as

$$C_{it} = \alpha C_{it-1} + X_{it}\beta + \in_i + J_{it} \quad , \qquad (4.6)$$

where C_{it} and C_{it-1} represent the contemporary and once-lagged crime rate respectively; X_{it} represents a vector of explanatory variables such as detection rate, prison population etc. Then, following Arellano and Bover (1995), each variable is taken "orthogonal deviation" to eliminate the time-invariant fixed effects $\in_i$ and thus equation (4.6) becomes

$$\tilde{C}_{it} = \alpha \tilde{C}_{it-1} + \tilde{X}_{it}\beta + \tilde{J}_{it} \quad . \qquad (4.7)$$

In equation (4.7), the variable $\tilde{C}_{it}$ is defined as

$$\tilde{C}_{it} = \left(C_{it} - \frac{1}{T_{it}}\sum_{s>t} C_{is}\right)\sqrt{\frac{T_{it}}{(T_{it}+1)}} \quad , \qquad (4.8)$$

[24] For a detail discussion please refer to Baltagi 2004; Hsiao 2003.

where T_{it} is the number of observations for each i. Equation (4.8) is also known as "forward orthogonal deviation" because it subtracts the contemporaneous observation from the average of all the future available observations.[25] Other variables in equation (4.7) are defined similarly.

By construction, $\tilde{C}_{it21}$ and $\mathcal{J}_{it}$ are correlated. Furthermore, if X_{it} is not strictly exogenous, there will also be correlation between $\tilde{X}_{it}$ and $\mathcal{J}_{it}$. For example, if X_{it} is detection rate or prison population which is not strictly exogenous, the forward orthogonal deviation of detection rate of prison population will be correlated with $\mathcal{J}_{it}$ in equation (4.7). Thus, to break the correlation between the error term and right-hand side variables in equation (4.7), instrumental variable estimation is applied. Potentially, all available lags of C_{it}, starting from C_{it22}, can be used as instruments for $\tilde{C}_{it21}$. In addition, all available lags of detection rate and prison population, starting from the second lag, are used as instruments for the forward orthogonal deviations of detection rate and prison population since they are also correlated with $\mathcal{J}_{it}$.[26]

Thus, the model specified in equation (4.5) is estimated with the GMM technique using the moment conditions generated by applying appropriate instruments. A robust variance-covariance matrix is used to obtain autocorrelation and heteroskedasticity corrected standard errors for the coefficients. The validity of the GMM estimation critically depends on the fact that the instruments are exogenous. As there are more

[25] Another way of removing the fixed effects $\in_i$ is to take the first difference instead of the orthogonal deviation, as proposed by Arellano and Bond (1991). However, unlike the orthogonal deviation approach, the first difference approach introduces first order serial correlation in the first difference regression errors.
[26] We are assuming that the detection rate and prison population are not strictly exogenous, rather they are predetermined.

instruments than the parameters to be estimated (over-identified model) in the extended model, the Sargan/Hansan test of over-identifying restrictions can be carried out to test for the validity of the instrument set.[27]

A potential problem with the "orthogonal deviation GMM" estimators is that moment conditions can increase prolifically. In particular, the number of instruments is quadratic in the time dimension of the panel. This can cause problems in finite samples. Specifically, since the number of elements in the estimated variance matrix of the sample moments is quadratic in the instrument count, it is quartic in T. A finite sample may lack adequate information to estimate such a matrix well (Roodman 2009). Furthermore, it can potentially weaken the Sargan/Hansen test for over-identifying restrictions to the point where it generates "very" good p-values of 1 (Anderson and Sorenson 1996).

Another point worth mentioning is that the asymptotics of this type of dynamic panel models are based on the "large N and small T" assumption. That is, the asymptotic properties, such as consistency, of the estimators are based on the assumption of large cross sectional dimension and small time dimension. The cross sectional dimension of 43 police force areas is not exactly "large" by normal standards. Hence, the results estimated by the GMM method in this chapter should be interpreted in light of this data limitation.

4.4 DATA DESCRIPTION

[27] Baltagi (2004) contains a detailed description of this test.

The empirical analyses in this chapter are based on two panel data sets in England and Wales. The first data set is used to provide the main estimation results and is disaggregated by 43 police force areas[28] in England and Wales covering the period 1992-2005. In this data set, prison population and Gini coefficient are aggregated on national level due to the data availability. All the other variables, on the other hand, are disaggregated by police force areas. The second data set is used to check the robustness of the results generated by the first data set, and it is disaggregated by 43 police force areas in England and Wales over the period 1987-2005. In the second data set, the prison population and Gini coefficient are aggregated on national level; the unemployment rate and real average weekly earnings are aggregated on regional level;[29] and the crime rate, detection rate and young people proportion are aggregated on police force area level. Therefore, the difference between the two data sets is that the first one provides more accurate measurements for the independent variables across areas but covers a shorter time period, while the second one is less accurate to measure the area-variations of independent variables but over a longer time period.

Table 4-1

Aggregation level of independent variables

Independent variables	1992-2005	1987-2005
Detection rate	Police force area	Police force area
Prison population	National	National
Gini coefficient	National	National
Proportion of young people	Police force area	Police force area

[28] There are in total 43 police force areas in England and Wales. They are Avon and Somerset, Bedfordshire, Cambridgeshire, Cheshire, Cleveland, Cumbria, Derbyshire, Devon and Cornwall, Dorset, Durham, Essex, Gloucestershire, Greater Manchester, Hampshire, Hertfordshire, Humberside, Kent, Lancashire, Leicestershire, Lincolnshire, City of London, Merseyside, Metropolitan Police District, Norfolk, Northamptonshire, Northumbria, North Yorkshire, Nottinghamshire, South Yorkshire, Staffordshire, Suffolk, Surrey, Sussex, Thames Valley, Warwickshire, West Mercia, West Midlands, West Yorkshire, Wiltshire, Dyfed-Powys, Gwent, North Wales, and South Wales.

[29] There are in total 10 government regions in England and Wales. They are North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East, London, South East, South West, and Wales (See Appendix II for regions' corresponding police forces).

Unemployment rate	Police force area	Regional
Real earnings	Police force area	Regional

4.4.1 Dependent Variable

The dependent variables in this chapter are six types of crime rates defined as the number of offences per 100,000 population. The crime rates being examined here are violence against the person, sexual offences, robbery, burglary, theft and handling, and fraud and forgery.[30] While the first three types are defined as violent crimes, the last three types are regarded as property crime. The crime rates data are obtained from two sources: (1) before the year 2001, they are collected from the annual command papers *Criminal Statistics* and (2) since 2001, they are obtained from the *Crime in England and Wales*.[31] Both these documents are published by the Home Office.

Two things about the crime rate data are worth mentioning. First, prior to 2001, the crime rates were measured as the number of offences recorded by police per 100,000 population. However, since 2001 the crime rates have been measured by the number of offences per 1000 population. Therefore, to have consistent series, the post-2000 crime rates have been multiplied by 100. One problem with the re-scaled crime rates is that the crime rates for the post-2000 period are less accurate and exhibit few, even no, variations over time and areas. This is particularly true for the sexual offence and robbery. Due to their nature, crime rates of sexual offences and robbery are much lower than the other crime rates included. Before the year 2001, the crime rates of sexual offences were lower than 100 per 100,000 population and those of robbery

[30] Criminal damage and other offences have been excluded from the analysis because there are large number of samples missing from the time periods under test.

[31] The serial numbers of *Criminal Statistics* are Cm847, Cm1322, Cm1935, Cm2134, Cm2410, Cm2680, Cm3010, Cm3421, Cm3764, Cm4162, Cm4649, Cm5001 and Cm5312. *Crime in England and Wales, 2001/02, 2002/03, 2003/04, 2004/05 and 2005/06*

were lower than 200 per 100,000 population for most police force areas. In the year 2001, the crime rates of sexual offences have become approximately 1 per 1000 population for most police force areas and those of robbery have been approximated to 1 or 2 per 1000 population. Being multiplied by 100, it is not surprising to see that the crime rates of sexual offences have been 100 per 100,000 population and those of robbery have become 100 or 200 per 100,000 population for most police force areas. As this situation is quite similar for the following 4 years in the data sets, the post-2000 crime rates present fewer variations in the analysis. Furthermore, there are some 0s in the crime rates of sexual offences and robbery in the year 2001 and onward which is also due to the reduced accuracy. Those 0s represent that the number of offences is less than 1 per 1000 population in some areas. Although it is probably not the case, the crime rates of these areas have to be 0 per 100,000 population after being multiplied by 100.

Second, there have been changes in the counting rules for the crime rates since April 1, 1998. These new rules have brought two major adjustments. First, the crime rates and relevant statistics have been documented according to the financial year system, which starts from 1st of April and ends on 31st of March the following year, rather than normal calendar year. Second, the definitions of some types of crime have been broadened and thus their crime rates exhibited upward shifts since 1998.[32] The most affected types are violence against the person and fraud and forgery and their crime rates have shown obvious upward shifts. Therefore, a dummy variable is included in

[32] The category of "violence against person" has been broadened to include harassment, cruelty to or neglect of children, assault on a constable and common assault. The "sexual offenses" category has been added with soliciting or importuning by man. "Theft and handling" has included the new sub-categories of vehicle interference and tampering. "Fraud and forgery" has been expanded to include bankruptcy and insolvency offenses and vehicle/driver fraud. The definitions of "robbery" and "burglary" have not been affected by the change in counting rules.

the analyses aiming to capture the changes in the counting rules since 1998. The dummy variable takes the value of one for the post-rule change period and zero otherwise.

The statistic summaries for the crime rates are given in table 4-2 and 4-3. As there are two sets of panel data being employed, table 1 summarizes the crime rates statistics for the data set 1992-2005 and table 4-3 presents the statistics for the data set 1987-2005. The summary tables also report the number of observations for each type of crime rate. There should be 602 observations in the data set 1992-2005 (43 areas by 14 years) and 817 observations in the data set 1987-2005 (43 areas by 19 years) if no observation is missing.

Table 4-2
Statistics for crime rates 1992-2005

		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	11,955	1,075	81	104
Median	9,133	743	67	63
Maximum	158,500	13,039	807	1,192
Minimum	3,797	180	0	0
Std. Dev.	16,484	1,294	71	137
Observations	602	602	602	602

	Property crimes		
	Burglary	Theft and handling	Fraud and forgery
Mean	1,984	5,735	668
Median	1,603	3,996	349
Maximum	16,827	101,400	23,100
Minimum	400	1,300	101
Std. Dev.	1,481	10,615	1,999
Observations	602	602	602

Table 4-3
Statistics for crime rates 1987-2005

		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	11,503	896	77	93
Median	8,778	517	61	42
Maximum	158,500	13,039	807	1,192
Minimum	3,797	175	0	0
Std. Dev.	16,690	1,173	71	144
Observations	817	817	817	817

	Property crimes		
	Burglary	Theft and	Fraud and

		handling	forgery
Mean	1,995	5,761	611
Median	1,586	3,900	300
Maximum	19,815	101,400	23,100
Minimum	400	1,300	97
Std. Dev.	1,751	10,955	1,842
Observations	817	817	817

4.4.2 Independent Variables

A host of independent variables are included in the analyses as proxies for the benefits and costs of committing crimes. These include detection rate, prison population, Gini coefficient, unemployment rate, real average weekly earnings and proportion of young people in the population. They are disaggregated on different levels because of the data availability.

4.4.2.1 Detection Rate

The detection rate is measured by the proportion of recorded offences that have been "cleared up". The "cleared up" offences are referring to those cases in which the offenders have been identified and given caution, fined or charged by the police. Therefore, the detection rate is included in the analyses as proxy for the probability of apprehension. According to Becker (1968) and Ehrlich (1973), higher probability of apprehension will increase the expected punishment of committing crimes and thus reduce the expected returns of doing so. As a result, the crime rate will be lowered due to the deterrent effect of higher probability of apprehension. The data of detection rate is disaggregated by police force areas in both data sets and obtained from the annual command paper *Criminal Statistics* with serial numbers provided before. The basic statistics of crime-specific detection rates are presented in the table 4-4.[33]

Table 4-4
Statistics for detection rates 1992-2005

[33] In order to save space and to keep the flow of the text, only statistics of independent variables for 1992-2005 are reported here, as it is the principal data set. All tables of statistics for independent variables, 1987-2005, are listed in *Appendix III*.

		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	28	72	64	32
Median	27	76	68	30
Maximum	69	97	124	96
Minimum	14	24	21	10
Std. Dev.	8	15	22	13
Observations	602	602	602	602

		Property crimes		
		Burglary	Theft and handling	Fraud and forgery
Mean		19	23	46
Median		16	22	44
Maximum		56	54	98
Minimum		7	8	9
Std. Dev.		8	7	16
Observations		602	602	602

4.4.2.2 Prison Population

The prison population is included as independent variable in the analyses to measure the severity of punishment. As demonstrated in Becker (1968) and Ehrlich (1973), more severe punishment is expected to reduce crime rate by increasing the expected cost of committing crimes for potential offenders. Therefore, when more offenders are sentenced into prison, potential criminals will expect more severe punishment if commit crimes and thus reduce their criminal activities. In addition, more prison population could reduce crime rate through the incapacitation effect. As more offenders sentenced into prison, it will be temporally impossible for them to commit further crimes while serving in prison. Thus, the crime rate will be lower for a period of time.

The prison population is calculated as the number of offenders sentenced into prison divided by the total population. In other words, it measures the number of offenders in prison per 100,000 population. The data is obtained from the annual command paper *Prison Statistics England and Wales* prior to year 2003. Since 2003, the data has been documented in the *Offender Management Caseload statistics*, a publication of the

Home Office. Although this variable varies by crime type, it is only available on national level in both data sets. The basic statistics of prison population are summarised in table 4-5.

Table 4-5
Statistics for prison population 1992-2005

		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	94	21	9	12
Median	100	21	9	12
Maximum	116	28	12	16
Minimum	64	14	6	8
Std. Dev.	18	4	2	2
Observations	602	602	602	602
		Property crimes		
		Burglary	Theft and handling	Fraud and forgery
Mean		14	10	10
Median		16	11	11
Maximum		17	12	12
Minimum		9	7	7
Std. Dev.		3	1	1
Observations		602	602	602

As shown in both table 4-5, the statistics of prison population for theft and handling and fraud and forgery are exactly the same. This is because the prison population of these two types of crime have been summed up and reported as one in the statistical publications. Since there is no way to separate the prison population between the two types of crimes, the summed prison population is included in the estimations of both theft and handling and fraud and forgery.

4.4.2.3 Gini Coefficient

Gini coefficient is an economic indicator measuring the degree of income inequality. The variable of Gini coefficient included in this chapter measures the inequality of post-tax income and it is aggregated on national level on both data sets. The data is obtained from the website of the *National Statistics*.

An increase in income inequality is expected to positively affect the crime rates, particularly property crime rates, by decreasing the opportunity cost of committing crimes for those at the lower end of the income distribution. Furthermore, as the income inequality reflects the wealth distribution between the rich and the poor, large increase in the income inequality could motivate the poor people to carry out anti-social behaviours or even violence crimes for revenge or releasing the anger. Therefore, higher income inequality could also lead to increased violent crime rates. The statistics of Gini coefficient will be reported later along with the other independent variables.

4.4.2.4 Unemployment Rate

The variable of unemployment rate is defined as the ratio of unemployment benefits claimants to the number of people in the workforce. The original data is available on both regional level and local authority[34] level with the region-level data being available for a longer time period. On the other hand, the data on local authority level is only available from the year 1992. Therefore, the unemployment rate in the data set 1992-2005 is aggregated from local authorities up to police force areas while the unemployment rate in the data set 1987-2005 is only on regional level. The original data source is the website of nomis – official labour market statistics.[35]

An increase in the unemployment is expected to have ambiguous effect on crime rates according to relevant literatures. On the one hand, higher unemployment rate will reduce the labour market opportunities for potential offenders and thus increase their

[34] "Local authorities" are lower government level than government regions but similar to county level. Data by local authorities typically contain six metropolitan counties, 27 non-metropolitan counties, 56 unitary authorities and the region of London, which has 32 London boroughs and the city of London.

[35]See https://www.noisweb.co.uk.

incentives to commit crimes through the motivation effect. On the other hand, higher unemployment rate will reduce the opportunities for certain crime types and thus reduce occurrence of those crimes due to the opportunity effect. For this reason, the net effect of unemployment on crime rates will depend on each specific case.

4.4.2.5 Real Average Weekly Earnings

The variable of real average weekly earnings is measured by the deflated average weekly earnings for all industries. The original data is available on both regional level and local authority level. The region-level data is collected from the *Regional Trends*, a Home Office publication, while the local authority data is from *Annual Survey of Hours and Earnings.*[36] Like unemployment rate, the variable of real earnings in the data set 1992-2005 is aggregated from local authorities into police force areas while real earnings in the data set of 1987-2005 is only aggregated on regional level.

The variable of real average weekly earnings is expected to be negatively correlated with crime rates. As real earnings measures the payoffs of legal labour market, an increase in the real earnings will increase the opportunity cost of committing crimes and thus reduce people’s intension to involve in illegal activities. Moreover, higher real earnings may also imply more opportunities for property crimes. Consequently, the net effect of real earnings on crime rates could be ambiguous and depending on the type of crime being analysed.

4.4.2.6 Young People Proportion

The variable of young people proportion is defined as the ratio between the number of young people aged between 15 to 24 years old and the entire population. The original

[36] *Regional Trends* is a publication that summarized various aspects of the U.K. by government regions. It summarizes the “real average weekly earnings” from the *Annual Survey of Hours and Earnings* published by the *Office for National Statistics*.

data is available on local authority level and has been aggregated into police force areas according to the geographic boundaries in both data sets. The data source is the mid-year estimated population by age groups and sex obtained from the *National Statistics*.[37] The number of people aged between 15 and 24 has been calculated by aggregating two original age groups 15-19 and 20-24.

Younger people are shown to be more prone to commit crimes than their older counterparts. One key reason for such behaviour is that the opportunity cost of committing crime is much lower for the younger people than their older counterparts. First, the younger people have, on average, lower earnings than their older counterparts. Therefore, if caught, they have less to lose when it comes to foregone earnings. Second, the penalty associated with committing crimes is probably lower for the younger people than their older counterparts because they usually receive lenient punishment and if they are under 18, their criminal records will be sealed by the age of 18. As a result, the proportion of young people could be positively correlated with crime rates in later empirical analyses. On the other hand, however, young people could also be deterred to commit crimes by the fact that, if they do so and get caught, their future labour market opportunities could be negatively affected by their criminal records. Hence, the net effect of the proportion of young people on crime could be ambiguous.

Table 4-6 below presents the key statistics for Gini coefficient, unemployment rate, real average weekly earnings, and the proportion of young people.

Table 4-6
Statistics for other independent variables 1992-2005

	Gini	Unemployment	Real average	Young people

[37] See http://www.statistics.gov.uk.

	coefficient		weekly earnings	
Mean	38.06	4.21	159.01	12.22
Median	38.00	3.57	143.98	12.18
Maximum	40.40	11.52	574.29	15.07
Minimum	36.10	0.61	65.59	8.11
Std. Dev.	1.13	2.40	65.62	1.08
Observations	602	602	602	602

4.5 RESULTS

In this section, we will firstly report the results of property crimes as they are expected to be more responsive to the controlling variables, especially to the social-economic factors, than the violent crimes. Next, the results of violent crimes will be reported to provide some interesting implications on the correlations between violent crimes and the controlling variables, which is less obvious to observe.

Each type of crime rate is analysed by three estimation methods as mentioned in the sub-section of empirical model and estimation methodology: the OLS estimation without fixed effects and lagged crime rate, fixed-effects estimation without lagged crime rate, and the GMM estimation on the full empirical model specified by equation (5). Furthermore, as there are two data sets being employed, the estimation based on the data set 1992-2005 will be reported first as the main results. Then the estimations based on the data set 1987-2005 will be presented to check the robustness of the coefficients. All variables, both dependent and independent, are taken logarithm before carrying out regression analysis.

4.5.1 Burglary

The estimation results of burglary based on the data set 1992-2005 are presented in table 4-7.

Table 4-7
Burglary 1992-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	12.62* (2.47)	12.56*** (1.34)	-
Crime(t-1)	-	-	0.53*** (0.06)
Detection Rate	-0.25 (0.17)	-0.15*** (0.03)	-0.08*** (0.01)
Prison Population	0.77*** (0.12)	-0.06 (0.08)	-0.10*** (0.03)
Gini Coefficient	-2.98 (2.04)	-1.04*** (0.19)	-0.31 (0.21)
Young People	0.91 (0.65)	-0.17 (0.23)	-0.16 (0.12)
Unemployment	0.36*** (0.10)	-0.06 (0.06)	-0.03 (0.04)
Real Earnings	0.44 (0.27)	0.06*** (0.02)	0.07*** (0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.40	0.94	0.67
S.E. of Regression	0.39	0.12	0.09
J-Statistic	-	-	41.84
Over Identification Test	-	-	0.17

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

As shown in the first column of table 4-7, the detection rate of burglary has negative but insignificant coefficient in the basic OLS estimation. Once the area-specific fixed effects are eliminated, the coefficient of detection rate has become negative significant at 1% level in the cross-sectional fixed effect model. When estimating with the GMM technique, the detection rate still has negative and highly significant correlation with burglary as shown in column three. Whilst all three estimations have generated negative coefficient for the detection rate as predicted, the magnitude of the coefficient is decreasing from column one to column three. This decreasing pattern is explainable: as we eliminated the area-specific fixed effects using fixed-effect and the GMM estimations, smaller share of the variations in burglary is explained by the detection rate. In addition, as additional relevant variable are controlled for (i.e. once-lagged crime rate), the correlation between detection rate and burglary becomes more significant.

The coefficient of prison population is positive and significant in the standard OLS estimation, contrary to the expectation. As the area-specific fixed effects are explicitly eliminated, the coefficient of prison population becomes negative but insignificant as shown in the second column. Subsequently, when the once-lagged burglary rate enters the equation as explanatory variable as well as both detection rate and prison population are instrumented by their lagged values to control for their endogeneity, the prison population has shown negative and highly significant correlation with burglary as predicted.

Furthermore, we do not find significant correlations between burglary and Gini coefficient, the proportion of young people, and unemployment. Gini coefficient has negative and insignificant coefficient in the OLS estimation, which turns to significant once the fixed effects are eliminated in the fixed-effect estimation. After further controlling measures applied for the endogenous variables as well as lagged crime rate in the GMM estimation, its coefficient becomes insignificant again. While the proportion of young people constantly shows insignificant effect on burglary, the coefficient of unemployment moves from positive and significant in the OLS estimation to negative and insignificant once the previously mentioned estimation issues are controlled in the GMM model. The insignificant correlation between unemployment and burglary is not surprising due to the two offsetting effects, namely motivation and opportunity. Additionally, the negative sign for unemployment is indeed consistent with the feature of burglary. As such crimes mainly target private dwellings, higher unemployment rate may significantly reduce the opportunities for burglary as more people would stay at or near their homes. As a result, although the coefficient is insignificant, the unemployment does seem to reduce burglary through opportunity effect.

The variable of real earnings has constantly shown positive effect across estimations. When the relevant estimation issues have been controlled for, the real earnings has shown positive and highly significant correlation with burglary in both fixed effect and the GMM estimations. This result suggests that the variable of real earnings has picked up the opportunity effect on burglary and higher income will increase the burglary rate by providing more opportunities.

There are two points worth mentioning about the GMM estimation. First, the once-lagged burglary rate has indeed obtained positive and significant coefficient as predicted. The magnitude of 0.53 suggests a quite strong self-correlation in burglary over time. Secondly, the over-identification test has shown that the instruments applied for the lagged burglary rate, contemporary detection rate and prison population are indeed uncorrelated with the error term and thus valid as instruments.

The estimation results based on the data set 1987-2005 are reported in table 4-8. These analyses are carried out to check whether the previously estimated coefficients are significantly affected by the changes in the data set.

Table 4-8
Burglary 1987-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	-4.76* (2.60)	11.11*** (2.09)	-
Crime(t-1)	-	-	0.71*** (0.03)
Detection Rate	-0.36*** (0.13)	-0.21*** (0.05)	-0.14*** (0.01)

Prison Population	-0.21***	-0.54***	-0.47***
	(0.05)	(0.04)	(0.01)
Gini Coefficient	2.90***	0.26	1.03***
	(0.46)	(0.31)	(0.04)
Young People	0.75*	-1.08***	-0.38***
	(0.42)	(0.40)	(0.11)
Unemployment	0.63***	0.14***	-0.07***
	(0.09)	(0.04)	(0.01)
Real Earnings	0.02	0.04	-0.03***
	(0.06)	(0.02)	(0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.35	0.90	0.81
S.E. of Regression	0.42	0.16	0.10
J-Statistic	-	-	41.27
Over Identification Test	-	-	0.18

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

As shown in table 4-8, both detection rate and prison population have shown negative and highly significant (at 1% level) correlation with burglary. These results are consistent with those based on the data set 1992-2005 as well as with the expectations.

The effect of Gini coefficient becomes positive and significant for the data 1987-2005. While this finding has enhanced the previous insignificant correlation between burglary and Gini coefficient, it also agrees with the expectation that higher income inequality will increase crime rates. Moreover, the effects of both young people and unemployment on burglary have been reinforced by applying the longer data set. Specifically, the proportion of young people still has the same negative sign as before which becomes significant in this case. This result denotes that an increase in the proportion of young people will reduce the crime rate of burglary. Such finding may be explained by the fact that young people, on average, have lower income levels than older people. Hence, while an increase in the number of young people could potentially increase the number of motivated offenders, it could also greatly reduce the opportunities for property crimes. Furthermore, by applying the data set 1987-2005, the study has also confirmed the previous finding that higher unemployment tends to reduce burglary by generating a negative and significant coefficient for unemployment. Therefore, in our analyses, unemployment mainly picks up the opportunity effect on burglary.

The only variable affected by using longer data set is real earnings and its coefficient changes from positive and significant in previous analyses to negative and significant now. This result implies that, while this variables picks up its opportunity effect on burglary, it reflects its motivation effect using the longer data set. As we eliminate the period 1987-1991 from the data set 1987-2005 and re-estimate the model, we obtain the same results to the previous finding that real earnings is positively and significantly correlated with burglary. Thus, we argue that the switched sign of this variable is caused by the extra information given by the years 1987-1991 and we will explain this in later discussion.

Unsurprisingly, the GMM estimation based on the data 1987-2005 has given a positive and significant coefficient for the once-lagged burglary rate. This rather high self-correlation of 0.71 has reinforced the previous finding that the crime rate of burglary is highly affected by its lagged value and suggesting a quite persistent pattern. The over-identification test has shown that the instruments employed for lagged burglary rate, contemporary detection rate and contemporary prison population are indeed valid as instruments for being independent from the error term.

4.5.2 Theft and Handling

The analyses of theft and handling are conducted by the same procedure. The results based on the data set 1992-2005 are reported in table 4-9.

Table 4-9

Theft and handling 1992-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	20.93* (11.71)	11.94*** (1.06)	-
Crime(t-1)	-	-	0.62*** (0.04)
Detection Rate	-0.65*** (0.19)	-0.23*** (0.04)	-0.14*** (0.03)
Prison Population	0.37*** (0.10)	-0.23*** (0.09)	0.06* (0.03)
Gini Coefficient	-3.61 (2.71)	-0.61*** (0.12)	-0.23** (0.10)
Young People	-0.18 (1.17)	-0.01 (0.24)	0.16 (0.10)
Unemployment	0.19 (0.11)	-0.03 (0.04)	0.02 (0.01)
Real Earnings	0.48 (0.35)	0.04*** (0.01)	0.08*** (0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.25	0.97	0.66
S.E. of Regression	0.46	0.09	0.07
J-Statistic	-	-	42.24
Over Identification Test	-	-	0.19

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

As shown in table 4-9, the variable of detection rate has constantly exhibited negative and significant correlation with theft and handling across estimations. This result reflects the expectation that detection rate, as proxy for the probability of apprehension, should negatively affect crime rate (theft and handling in this case) through both deterrence and incapacitation effects. The prison population, on the other hand, has displayed unstable coefficient over estimations. Its effect is positive and significant in the OLS estimation, changes to negative and highly significant in the fixed effect model, and switches back to positive and significant at 10% level in the GMM model. Given the estimation result generated by the GMM method, we should be cautious to conclude that prison population is positively correlated with theft and handling because there may be several reasons for the positive effect of prison population. First, higher prison population may be caused by higher crime rate if the detection rate remains unchanged. When there are more offenders and the same proportion of them has been detected, more offenders will be sentenced into prison accordingly. Second, the variable of prison population is aggregated on national level and thus cannot reflect its variations by areas. Third, the prison population of theft and handling is reported together in data with that of fraud and forgery as one category. Hence, this limitation of data may also bias the estimation results.

The Gini coefficient has shown negative correlation with theft and handling in all three estimations, insignificant in the OLS estimation and significant in fixed-effect and the GMM estimations. These results are contrary to the expectation that higher income inequality will increase property crime rates through reducing the opportunity cost of committing crimes by the poor. However, as higher income inequality also indicates that larger proportion of wealth is possessed by smaller proportion of population, the opportunities for property crimes could be reduced consequently as fewer wealthy people will be around. Another possible explanation is that the variable of Gini coefficient is aggregated on national level and thus exhibits no variation across areas. For this reason, its coefficient may be biased by this data limitation.

Based on our analyses, we do not obtain significant effects of young people proportion and unemployment on theft, just like the case of burglary. However, this alone is not enough to conclude that there is no correlation between them and the crime of theft. This is because, although the proportion of young people is predicted to positively affect crime given that they are more crime-prone, we also believe that such positive effect could be counterbalanced by the fact that young people, on average, have lower incomes (if any) and possess less valuable goods and provide fewer opportunities for property crimes. Therefore, the estimated coefficient for the proportion of young people could reflect both effects and appears to be insignificant. Likewise, unemployment is also well-known for having two opposite effects on crimes: motivation and opportunity. Moreover, as higher unemployment often indicates an economy downturn, people are less likely to consume on expensive goods and thus provide fewer targets for potential thieves. As a result, the estimated insignificant coefficient potentially suggests that motivation and opportunity effects are roughly equally strong.

Our results suggest a positive correlation between real earnings and theft and such correlation is insignificant in the OLS estimation and significant at 1% level once we control for area-specific fixed effects and endogenous variables. This result implies that real earnings mainly picks up its opportunity effect on theft, just like the case of burglary: higher income levels will increase people's consumption on valuable goods and thus provide more opportunities for property crimes.

In addition, our GMM estimation shows a rather strong positive correlation between the contemporary and once-lagged crime rate of theft with a magnitude of 0.62. Furthermore, the over-identification test has captured an insignificant *p*-value of 0.19, indicating that the null hypothesis of valid over-identifying restrictions cannot be rejected and hence the employed instruments are truly valid.

We report our results using the data set 1987-2005 in table 4-10 which can be compared with previous results. However, it is important to note that, in the GMM estimation, being most suitable for our empirical model, we have attained highly significant *p*-value for the over-identification test. Such result implies that the instrument variables we adopt for the endogenous variables (lagged crime rate, contemporary detection rate and prison population) are invalid as instruments probably due to their correlations with the error terms. Consequently, the estimations generated by the GMM model lose their power to be unbiased and consistent, thus cannot be used as solid evidence.

Table 4-10
Theft and handling 1987-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	7.66 (5.31)	11.91*** (0.83)	-
Crime(t-1)	-	-	0.42*** (0.03)
Detection Rate	-0.96** (0.38)	-0.27*** (0.03)	-0.26*** (0.02)
Prison Population	-0.23* (0.13)	-0.52*** (0.03)	-0.45*** (0.01)
Gini Coefficient	1.19** (0.49)	-0.10 (0.17)	0.55*** (0.05)
Young People	-0.09 (1.00)	-0.45*** (0.08)	-0.41*** (0.04)
Unemployment	0.43*** (0.15)	0.12*** (0.03)	0.07*** (0.02)
Real Earnings	-0.09 (0.08)	-0.01 (0.01)	-0.07*** (0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.29	0.96	0.73
S.E. of Regression	0.45	0.10	0.08
J-Statistic	-	-	58.48
Over Identification Test	-	-	0.01

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

As we check the results, the analyses based on the longer data set have confirmed the effects of lagged crime rate, detection rate, and unemployment rate. In the meantime, they have provided opposite results for prison population, Gini coefficient and real earnings. However, as we shall remember that the results given by the GMM estimation are questionable due to the invalid instruments, such results cannot be used as solid evidence to generate further inferences.

4.5.3 Fraud and Forgery

As usual, we report the results based on the data set 1992-2005 as our main results in table 4-11 below.

Table 4-11

Fraud and forgery 1992-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	24.86* (13.27)	10.97*** (2.41)	-
Crime(t-1)	-	-	0.23*** (0.02)
Detection Rate	-0.46*** (0.08)	-0.30*** (0.07)	-0.45*** (0.06)
Prison Population	0.39* (0.23)	-0.23 (0.19)	-0.39*** (0.11)
Gini Coefficient	-5.62 (3.49)	-1.06*** (0.31)	-0.62*** (0.22)
Young People	-0.40 (1.15)	-0.01 (0.47)	-1.01*** (0.32)
Unemployment	-0.10 (0.14)	-0.13 (0.11)	-0.11* (0.06)
Real Earnings	0.78* (0.46)	0.12*** (0.03)	0.03* (0.02)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.34	0.91	0.73
S.E. of Regression	0.58	0.22	0.19
J-Statistic	-	-	32.06
Over Identification Test	-	-	0.56

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.

The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

According to table 4-11, we find that fraud and forgery is negatively affected by both detection rate and prison population at 1% significance level. Whilst the coefficient of detection rate is constantly negative and highly significant, the effect of prison population is less stable across estimations. It is positive and significant at 10% level in the OLS estimation, negative and insignificant in the fixed-effect model, and becomes negative and highly significant (at 1% level) in the GMM model. These results support the expected negative effects of both detection rate and prison population as law enforcement instruments.

Same as for burglary and theft, we achieve negative correlation between Gini coefficient and fraud and forgery. Since the variable of income inequality has constantly shown negative effect on property crime rates that have been analysed, it may be argued that such negative effect is due to the fact that, with higher income inequality, larger proportion of wealth is shared by fewer people implying fewer opportunities for property crimes. Moreover, both unemployment and the proportion of young people have negative effects on fraud and forgery under the GMM estimation. We may explain the negative effect of young people with the fact that over 90 percent of the cases in fraud and forgery are cheque and credit card fraud and, hence, the opportunities of this crime should be negatively correlated with the number of young people. This is because, in general, they have lower income and should spend less both in store and online. For this reason, the predicted positive correlation between the proportion of young people and crime rate could be offset by the reduced crime opportunities due to higher proportion of young people. Consequently, the net effect of young people proportion on property crime rates, fraud and forgery in this case, may be insignificant or even negative. Furthermore, unemployment rate has also picked up its opportunity effect by showing a negative and significant coefficient. This result can be again explained by the fact the majority of this type of crime are cheque and credit card fraud; an increase in unemployment will not only reduce people's spending by cheques and credit cards, but also reduce the opportunities of potential offenders to commit such crimes as they normally need jobs to do so. In addition, our results suggest that real earnings is positively correlated with fraud in all cases implying that an increase in earnings will lead to higher rate of fraud. Unsurprisingly, this variable has, once again, picked up its opportunity effect just as it did in the analyses of both burglary and theft.

The once-lagged crime rate shows a positive and significant (at 1% level) correlation with the contemporary crime rate with the magnitude of 0.23. This result confirms our prediction that crime rate should be persistent over time. Also, the over-identification test has revealed a *p*-value of 0.56 indicating the instrument variables that we implemented for lagged crime rate, contemporary detection rate and prison population are truly valid and have supplied with useful information for estimating our model.

For the purpose of comparison, we also report the results based on the data set 1987-2005 are in table 4-12.

Table 4-12
Fraud and forgery 1987-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	9.97* (5.51)	10.78*** (2.14)	-
Crime(t-1)	-	-	0.30*** (0.08)
Detection Rate	-0.55*** (0.13)	-0.32*** (0.06)	-0.28*** (0.07)
Prison Population	-0.36** (0.10)	-0.46*** (0.09)	-0.40*** (0.06)
Gini Coefficient	-0.23 (0.71)	-0.52 (0.41)	-0.20 (0.19)
Young People	-0.39 (1.06)	-0.38* (0.21)	-0.44 (0.38)
Unemployment	0.10 (0.14)	0.01 (0.08)	0.05 (0.04)
Real Earnings	0.15** (0.05)	0.04 (0.03)	-0.03** (0.02)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.25	0.89	0.71
S.E. of Regression	0.63	0.24	0.19
J-Statistic	-	-	37.44

Over Identification Test	-	-	0.31

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

Generally speaking, our analyses based on the longer data set have generated reasonably consistent results as our previous findings. Specifically, both detection rate and prison population are negatively correlated with fraud and forgery and the correlations are significant at 1% level most of the time. Moreover, Gini coefficient and the proportion of young people still have negative effects on fraud reflecting their opportunity effect on property crime, although their coefficients become less significant.

The only significant change made by applying the longer data set is the effect of real earnings on fraud: while the effect is positive earlier reflecting the opportunity effect, it becomes negative and significant when using the longer data set picking up its motivation effect. As we have observed the same change in the analyses of burglary, we will explain such phenomenon together in later discussion.

The GMM estimation shows that the once-lagged crime rate is correlated with contemporary crime rate at a rate of 0.30, which is coherent with the previous finding. Furthermore, the over-identification test has also confirmed the validation of employed instruments.

4.5.4 Robbery

According to the Home Office, robbery has been categorized as violent crime because the occurring of such crimes is usually accompanied with physical harm implied onto the victims. However, it cannot neglect that the motivation of committing such crimes is usually of financial interest. Therefore, robbery is expected to respond in similar way to the independent variables as do with property crimes. The estimation results based on the primary data set 1992-2005 are reported as following in table 4-13.

Table 4-13

Robbery 1992-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	116.61 (615.40)	77.59 (132.60)	-
Crime(t-1)	-	-	0.44*** (0.004)
Detection Rate	-3.59*** (1.26)	-0.88** (0.48)	-3.42*** (0.03)
Prison Population	4.63 (4.75)	-0.04 (1.72)	-2.59*** (0.32)
Gini Coefficient	-6.74 (12.18)	-0.62 (2.28)	1.03** (0.50)
Young People	17.36* (10.09)	16.10 (10.28)	1.03* (0.60)
Unemployment	2.52 (12.44)	-21.02** (8.52)	-12.69*** (0.66)
Real Earnings	0.80** (0.40)	0.15*** (0.04)	0.09*** (0.003)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.28	0.89	0.33
S.E. of Regression	115.72	45.02	35.70
J-Statistic	-	-	36.23
Over Identification Test	-	-	0.36

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

The first thing to notice in table 4-13 is that both detection rate and prison population have displayed negative and significant impacts on robbery as law enforcement instruments. Whilst the coefficient of detection rate is constantly negative and significant in all three estimations, the effect of prison population is insignificant in both the OLS and fixed-effect estimations and becomes highly significant (at 1% level) once the relevant estimation issues are controlled by the GMM technique.

Meanwhile, we also find that both Gini coefficient and the proportion of young people present positive and significant correlations with robbery, thus supporting our expectations. As we have predicted that an increase in the income inequality will increase crimes by reducing the opportunity cost of committing crimes for the people at the lower end of wealth distribution, our results have indeed reinforced such predictions. Furthermore, the proportion of young people could positively affect robbery through two channels: on the one hand, higher proportion of young people increases the number of motivated robbers given the lower opportunity cost of young population; on the other hand, higher proportion of young people also means more potential targets for robbery. Due to its nature, robbery often occurs on the streets of less affluent areas. In addition, the targets of robbery are usually the valuable belongings carried by passengers and pedestrians, such as mobile phones, ipods, laptops, wallets and purses and so on. Consequently, for robbers, young people are usually seen as attractive targets since they are relatively less cautious and more likely to possess trendy electronic gargets. As a result, higher proportion of young people increased the number of both potential robbers and potential victims, hence, significantly increases the crime rate of robbery.

Whilst unemployment rate shows negative and significant correlation with robbery, real earnings exhibits positive and significant effect on the same crime. Both of them have picked up their opportunity effects. In particular, as long as the potential estimation issues are controlled, higher unemployment rate will reduce the crime rate of robbery, reflecting the fact that robbery primarily occurs on the streets and, the more people being unemployed, the fewer commutes would be required. Thus, it provides fewer opportunities for potential robbers. Whereas, with higher income, people tend to make more purchases, thus provide more opportunities for robbery.

The coefficient of once-lagged robbery rate is positive and significant at 1% level. The magnitude of 0.44 suggests quite strong persistence in robbery which is again consistent with expectation. Meanwhile, the over-identification test has confirmed that the instruments employed for the endogenous variables in the GMM estimation are indeed valid and independent from the error term.

Table 4-14 below shows the counterpart results generated by the data set 1987-2005.

Table 4-14

Robbery 1987-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	-662.20*** (254.37)	-215.79 (258.73)	-
Crime(t-1)	-	-	0.76*** (0.001)
Detection Rate	-4.14*** (1.24)	-0.41 (0.44)	-2.18*** (0.04)
Prison Population	17.96*** (5.68)	-2.00 (2.55)	-8.26*** (0.93)
Gini Coefficient	10.64*** (4.02)	1.83 (2.84)	0.85* (0.47)
Young People	18.60* (10.80)	15.73 (12.59)	-1.67*** (0.39)
Unemployment	12.22** (5.14)	0.75 (1.72)	-3.15*** (0.54)
Real Earnings	0.40** (0.18)	0.16*** (0.04)	0.06*** (0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.24	0.87	0.59
S.E. of Regression	123.63	51.97	35.08
J-Statistic	-	-	34.55
Over Identification Test	-	-	0.49

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.

The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

As shown in table 4-14, the estimation results based on the data set 1987-2005 are quite consistent with previously findings, confirming the robustness of our results. Especially, both detection rate and prison population have shown negative correlations with robbery indicating their deterrent (and probably also incapacitation) effects on crime as law enforcements. Gini coefficient displays positive and significant correlation with robbery implying that higher income inequality will increase robbery through lowering the opportunity cost to the least affluent. Additionally, both unemployment rate and real earnings have picked up their opportunities effects on robbery: higher unemployment rate will reduce the opportunities and hence the crime rate for robbery; while higher income will increase the opportunities for potential robbers and hence increase the crime rate of robbery.

The only exception is the proportion of young people which has obtained unstable coefficients across estimations. The correlation between the proportion of young people and robbery is positive in both the OLS and fixed-effect estimations. Once the GMM estimation is applied, the coefficient of young people turns into negative and significant at 1% level.[38] Therefore, the coefficient changing is probably as a result of the inclusion of extra 5 years in the data set 1987-2005. In this case, the coefficient of young people is rather sensitive to the period under examination.

The GMM estimation based on the data set 1987-2005 has confirmed the significant self-correlation in robbery. In addition, the over-identification test has proved the validity of using lagged values as instruments for the endogenous variables.

[38] By restricting the data set 1987-2005 to only include the period 1992-2005, the coefficient of young people are constantly positive, which is consistent with the previous finding.

4.5.5 Sexual Offences

Sexual offence is a typical violent crime and does not pursue financial benefit. Such crime may also be affected by the explanatory variables, but probably in a less obvious way. For example, the law enforcement variable may still have deterrence or incapacitation (or both) effect on sexual offensive activities. Higher unemployment rate may reduce the potential contacts between potential offenders and victims, so it reduces the opportunities of such crime. Similarly, higher incomes could also reduce the possible contacts between offenders and victims. This is because, with higher income, the potential victims (mostly women), are able to afford more self-protections, such as cars, instead of walking or taking public transportation. For these reasons, sexual offences has also been analysed in the same empirical framework as property crimes in hope to reveal some less obvious relationships between this type of crime and potentially relevant factors. The analyses using the data set 1992-2005 are reported below in table 4-15.

Table 4-15

Sexual offences 1992-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	737.29* (437.49)	69.90 (76.62)	-
Crime(t-1)	-	-	0.22*** (0.001)
Detection Rate	-1.04* (0.56)	0.20 (0.40)	1.37*** (0.05)
Prison Population	-1.88 (4.13)	-9.95*** (2.53)	-15.25*** (0.71)
Gini Coefficient	-12.65* (7.47)	-1.98 (1.54)	-1.44*** (0.33)
Young People	-10.36 (9.65)	4.79** (2.07)	0.51 (0.55)
Unemployment	-1.42 (3.94)	3.39 (2.32)	9.84*** (0.33)
Real Earnings	0.42* (0.25)	0.03* (0.02)	0.08*** (0.003)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.25	0.81	0.03
S.E. of Regression	61.67	31.40	32.15
J-Statistic	-	-	37.39
Over Identification Test	-	-	0.32

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.

The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

The detection rate of sexual offences has changing coefficient across estimations. In the OLS estimation, the coefficient is negative and significant (at 10% level), confirming the expected deterrent effect of law enforcement. As the area-specific fixed effects are eliminated in the second estimation, the coefficient turns positive and insignificant. Once the GMM technique is applied to further control the possible estimation issues, detection rate has shown positive and highly significant (at 1% level) correlation with sexual offences. The result of GMM estimation, although contrary to expectation, is reasonable due to the seriousness of sexual offences. As this type of crime is so severe that it induces enormous damage to both the victims and the society, higher rate of sexual offences would demand higher probability of detection to combat the crime. As a result, although we have controlled the endogeneity of detection rate, its positive coefficient could still reflect the positive correlation between sexual offences and detection rate.

The prison population has attained overall negative coefficient, which is insignificant in the OLS estimation but significant in both fixed-effect and the GMM estimations. This result is in line with the expectation and probably reflects the incapacitation effect of prison population. The motivation of such offences is obviously not for financial gains, but it should be rather heavily correlated with the offenders' psychological characteristics. As more offenders are jailed and temporally separated from the potential victims, there will be less offenders remaining in the society. This eventually would reduce reoccurrences of sexual offences.

Gini coefficient shows negative correlation with sexual offences in all estimations. Its coefficient is significant at 10% level in the OLS estimation, insignificant in the fixed-effect estimation, and highly significant (at 1% level) in the GMM estimation. However, it is still not enough to conclude that income inequality, as a macro-economic indicator, is negatively related with sexual offences. Hence, further investigation in their correlation, if any, is necessary. The coefficient of young people is negative and insignificant in the OLS estimation. It becomes positive and significant in the fixed-effect model and changes to positive and insignificant in the GMM model. Given the results, there seems no strong evidence indicating a causal relationship between the proportion of young people and sexual offences. Furthermore, we find that both unemployment and real earnings have achieved positive coefficients in most cases. Nonetheless, we still cannot conclude with confidence that an increase in either unemployment or income will lead to higher crime rate of sexual offences. This type of crime deserves formal analysis that specially designed to suit its characteristics.

One of the results that do make sense is the significant self-correlation in sexual offences at a rate of 0.22 as suggested by our GMM estimation. This result is consistent with the expectation that crime rate is persistent. In this case, the offenders commit such crimes mainly due to their different “taste” from others. Therefore, as long as they are not detected and sentenced into prison, they will probably keep committing such crimes because their “taste” does not disappear.

We use table 4-16 to present the results estimated from the data set 1987-2005.

Table 4-16

Sexual offences 1987-2005

variable	OLS	Cross-Section	GMM

		Fixed Effect	
C	140.21 (88.33)	56.88 (86.63)	-
Crime(t-1)	-	-	0.24*** (0.002)
Detection Rate	-1.64* (0.98)	0.10 (0.37)	0.38*** (0.01)
Prison Population	54.54 (47.83)	-10.09** (4.05)	-13.68*** (0.64)
Gini Coefficient	-2.76 (1.69)	-2.54* (1.42)	-1.51*** (0.20)
Young People	-9.40 (11.41)	9.58*** (2.06)	9.26*** (0.17)
Unemployment	6.38 (5.65)	-1.95*** (0.53)	-2.39*** (0.12)
Real Earnings	0.18 (0.13)	0.01 (0.02)	0.02*** (0.004)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.16	0.82	0.04
S.E. of Regression	65.62	30.51	29.92
J-Statistic	-	-	41.21
Over Identification Test	-	-	0.22

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

We have found that estimating the empirical model with the data set 1987-2005 has produced largely consistent results as our previous findings. First, the coefficient of detection rate is negative and significant at 10% level in the OLS estimation, positive and insignificant in the fixed-effect estimation, positive and highly significant in the GMM estimation. The changing pattern of the coefficient is exactly the same as previous results and has confirmed that sexual offences is indeed positively correlated with its detection rate. Second, prison population has shown negative and significant impact in both fixed-effect and the GMM estimations. This result is also consistent with the previous finding that sexual offences is negatively affected by the imprisonment rate. Third, the variables of Gini coefficient, young people and real earnings have also displayed the same signs as before. However, as mentioned earlier, we are unable to explain their estimated impacts on sexual offences with confidence due to the lack of established theories that relate these variables to sexual offences. In addition, the GMM model has also confirmed the positive self-correlation in the crime rate of sexual offences, as well as the validity of applied instruments for the endogenous variables.

The only exception is the unemployment rate. In the analyses based on the longer data set 1987-2005, the coefficient of unemployment is positive and insignificant in the OLS estimation, negative and highly significant (at 1% level) in both fixed-effect and the GMM estimations. This result supports the expectation and advocats that higher unemployment will reduce the crime rate of sexual offences probably through reducing the opportunities for such crimes.

4.5.6 Violence against the Person

Violence against the person refers to those offences which have harmed or potentially harmful to the physical well-beings of individuals. In some cases, however, the offenders carried out such violent actions still aiming for financial gains. Under such circumstances, the offences will be recorded as violence against the person as long as the victims are physically harmed or threatened. The crime rate of violence against the person could be, at some level, correlated with the economic variables in a similar way to property crimes. The empirical analyses are firstly conducted using the data set 1992-2005 as usual, the results of which are reported in table 4-17.

Table 4-17

Violence against the person 1992-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	20.03** (8.78)	12.56*** (2.90)	-
Crime(t-1)	-	-	0.41*** (0.03)
Detection Rate	-0.82** (0.42)	-0.24* (0.12)	-0.51*** (0.07)
Prison Population	0.12 (0.67)	-0.58 (0.42)	1.39*** (0.18)
Gini Coefficient	-3.28* (1.87)	-1.88*** (0.53)	-2.61*** (0.19)
Young People	-0.32	0.91**	0.16

	(1.02)	(0.42)	(0.10)
Unemployment	0.23**	0.002	-0.06
	(0.13)	(0.10)	(0.06)
Real Earnings	0.31	0.06***	0.16***
	(0.24)	(0.02)	(0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.69	0.92	0.73
S.E. of Regression	0.42	0.21	0.20
J-Statistic	-	-	40.18
Over Identification Test	-	-	0.25

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

As the detection rate has constantly shown negative and significant impact on the personal violence in all three estimations, prison population has attained positive correlation with such crime in the GMM estimation. Whilst we acquire negative and significant effect of Gini coefficient on personal violence, the proportion of young people is positively, but insignificantly, correlated with this crime rate. Moreover, whilst unemployment rate shows negative and insignificant effect on violent crimes, real earnings is positively correlated with such crimes. Given the fact that some personal violence are committed for economic returns as mentioned earlier, one potential explanation for both coefficients of unemployment and real earnings could be that either higher unemployment or lower income may reduce personal violence through their opportunity effects.

The coefficient of lagged crime rate suggests that violence against the person, like other crimes, is positively self-correlated. Meanwhile, the over-identification test cannot reject the validity of the employed instruments by generating an insignificant *p*-value of 0.25.

The estimation results based on the data set 1987-2005 are presented in table 4-18.

Table 4-18

Violence against the person 1987-2005

variable	OLS	Cross-Section Fixed Effect	GMM
C	6.82** (3.37)	8.53*** (2.11)	-
Crime(t-1)	-	-	0.39*** (0.04)
Detection Rate	-0.94** (0.47)	-0.28** (0.11)	-0.33*** (0.08)
Prison Population	0.22 (0.29)	-0.25*** (0.09)	-0.01 (0.03)
Gini Coefficient	0.39 (0.45)	-1.05*** (0.40)	-1.12*** (0.13)
Young People	0.22 (0.94)	1.09*** (0.27)	0.11 (0.09)
Unemployment	0.32*** (0.11)	-0.10 (0.07)	-0.08** (0.04)
Real Earnings	-0.08** (0.04)	-0.003 (0.02)	0.02 (0.01)
Dummy	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes
Adjusted R Squared	0.71	0.93	0.73
S.E. of Regression	0.42	0.21	0.18
J-Statistic	-	-	74.53
Over Identification Test	-	-	0.0001

*** significant at 1% level; ** significant at 5% level; * significant at 10% level.
The J-statistic is computed for the Sargan/Hansen over-identifying restrictions. Under the null hypothesis that the over-identifying restrictions are valid, the J-statistic follows a Chi-Squared distribution with the degree of freedom being the difference between the instrument rank and the number of coefficients estimated. The reported over-identification test is the corresponding *p*-value.

The results generated by applying the data set 1987-2005 have proved our previous findings are rather robust. In particular, we find negative and significant coefficient for the detection rate, whilst prison population has negative but insignificant correlation with personal violence. Furthermore, Gini coefficient still shows strong negative impact on personal violence implying that higher income inequality will reduce such crime rate. The proportion of young people has displayed positive coefficient which is also consistent with the previous result. In addition, while unemployment rate mainly demonstrates negative effect, real earnings is positively, but insignificantly, correlated with violent crime.

With the longer data set, the once-lagged crime rate obtains a coefficient of 0.39, suggesting a significant self-correlation over time. Nonetheless, given the *p*-value of 0.001, the over-identification test has rejected the validity of using the lags of endogenous variables as instruments.

4.6 CONCLUSION

In this chapter, we have performed various panel data analyses to indentify the effects of crime-influential factors on different types of crime rates in England and Wales. We are motivated by the fact that literature in England and Wales investigating similar problems are limited in both aspects of quantity and quality, especially when comparing to literature in the United States. As we have surveyed the existing literature that detecting the determinants of crimes in England and Wales using panel data, we have found some potential issues that may bias the results of those literature, such as uncontrolled endogeneity of law enforcement variables, omitted variables and so on. Therefore, we tackled those associated estimation issues in the work of this chapter by adopting appropriate methodologies and variables.

The crime rates of our interest are violence against the person, sexual offences, robbery, burglary, theft and handling, and fraud and forgery. We choose the explanatory variables strictly according to Becker (1968) and Ehrlich (1973) to reduce the risk of omitting relevant variables and to incorporate the factors representing law enforcement, economic conditions, as well as demographic feature. In addition, our panel data aggregated by police force areas has enabled us to avoid the potential bias caused by the correlation between independent variables and area-specific fixed effects. As a further temptation to include a relatively complete set of explanatory variables, we also incorporate once-lagged crime rate into the right-hand side of our empirical model as we believe it may measure the persistence in crime as well as the effects of lagged independent variables. To cope with our model specification as well as the endogeneity of law enforcement variables, we have applied the GMM technique to produce unbiased estimates. We report our main results based on the panel data set cover the period 1992-2005. However, as a robust check, we also present the results based on the data set 1987-2005 afterwards. The main difference between the two data sets is the aggregation level of two independent variables: while unemployment and real earnings are aggregated by police force areas in the data set 1992-2005, they are on regional level in the data set 1987-2005.

We can summarise the estimation results as follows. First, the property crimes are better explained by our empirical model. While analysing property crimes, the estimated coefficients for independent variables are more consistent with expectations than violent crimes and easier to explain in the case of they are not supporting the expectations.

Second, detection rate and prison population are the strongest predictor for both violent and property crimes. They have shown negative and significant effects on different crime rates in most cases confirming their effectiveness as law enforcement instruments. This result does not get affected by applying different data sets and is very consistent across different estimations. The only exceptions are that prison population is positively correlated with theft and handling and violence against the person whilst the detection rate is positively correlated with sexual offences. Since it is possible that more law enforcement effort is required by higher crime rates, the positive correlations between law enforcement instruments and certain types of crime rates are explainable in this way.

Third, the social-economic factors of unemployment rate and real earnings, which are also our main concern, have also obtained stable and reasonable correlations with our examined crime rates, particularly property crime rates. Our results show that they have both picked up their opportunity effects on the crime rates of burglary, theft and handling, fraud and forgery as well as robbery and these findings are not affected by using longer data set 1987-2005 in most cases. However, we do find two exceptions: the variable of real earnings shows its opportunity effect on burglary and fraud based on the data 1992-2005. Once we apply the longer data set, real earnings picks up its motivation effect for both burglary and fraud. As we delete the 5 extra years (1987-1991) from the data set 1987-2005, the estimation results confirms the previous findings that real earnings mainly shows its opportunity effects. Hence, we are able to conclude that by including the period 1987-1991 into the sample, it would alter the effect of this variable in the analyses of burglary and fraud.

The difference in the coefficient of real earnings as discussed above may be explained by macroeconomic situation at the early 1990s. The UK experienced a period of serious recession between 1990 and 1992, triggered by saving and loan crisis in the United States. During this recession, important economic indicators such as real earnings also experienced deterioration. The data set 1992-2005 principally measures the correlation between crime rates and social-economic factors in a period of economic growth. As demonstrated in our empirical analysis, during this period, the social-economic factors mainly picked up the opportunity effects. This is because on the one hand, people are more likely to spend under good economic conditions, hence, providing relatively more opportunities for property crimes. On the other hand, potential offenders may not be facing serve economic difficulties in times of prosperity. Property crimes committed may largely due to increasing opportunities, not economic desperations. Consequently, during economic growth, the opportunity effects appear to be stronger than the motivation effects. On the contrary, the data set 1987-2005 has incorporated the period of economic recession in the early 1990s. During economic uncertainties, the correlations between social-economic factors and property crimes should be dominated by the motivation effects, rather than the opportunity effects. This is because, in recessions, potential offenders are more likely to be facing financial problems, which may drive them to commit property crimes as a solution. An inevitable decrease in the earnings during economic downturn, not only create fewer opportunities for property crimes, as people tend to spend less, more importantly, it also may turn more people into potential criminals due to their economic desperations. As a result, the motivation effects appear to be stronger than the opportunity effects, in times of economic crisis. In summary, adding the period 1987-1991 into the data set has altered the net effects of real earnings on burglary and fraud.

Another potential reason for the change in the coefficient of real earnings is that John Major became the leader of the Conservatives in 1990. Therefore, while the period 1992-2005 is only under the leaderships of John Major and Tony Blair, the period of 1987-2005 is also governed by Margaret Thatcher who was in office from 1979 to 1990. While Margaret Thatcher focused on social and economic reform which led to relatively unstable environment, John Major is famous for his mild-mannered governing style. Under his leadership, the economy in the UK stably grew over time and the unemployment rate was back in control and stayed on normal level. Such positive atmosphere could change people's style of consuming and saving and, therefore, the opportunities and motivations of illegal activities.

Fourthly, the variable of Gini coefficient has constantly exhibited negative and significant correlation with different crime rates, which is opposite to expectation. Apart from the fact that Gini coefficient is aggregated on national level and cannot well reflect the variations of income inequality across areas, its negative effect may be explained by the reason that higher income inequality will reduce the number of rich people and thus reduce the crime opportunities. The proportion of young people shows mixed influences on different crime rates and is insignificant in most cases. However, we explain the unstable impact of young people by arguing that, whilst higher proportion of young people implies more motivated offenders, it could also reduce the opportunities for crimes, particular property crimes, because they are, on average, having lower income levels and processing less valuable goods. Also, as we argues before, the severe consequence of committing crimes and getting caught for young people, such as less labour market opportunities in the future, could prevent them from participating in illegal activities and thus offset their positive effects on crime.

The fifth conclusion can be drawn from this chapter's analyse is that, consisting with expectation, each type of crime rate is significantly correlated with its once-lagged value suggesting strong persistence over time. This may be due to the offenders' persistent preference for crimes, the impacts of lagged independent variables, as well as the impact of lagged business cycle.

Whilst we have attempted to provide unbiased results by using the most refined data set and appropriate estimation methodologies available, we shall nevertheless acknowledge the potential limitations of this work. First of all, the prison population and Gini coefficient are aggregated on national level and shows no variation across areas. Such data limitation could affect the estimation results. Second, the empirical model adopted in this chapter is unable to distinguish the motivation effect from the opportunity effect for some independent variables. As predicted by the theories, several variables could have the double-edged effect on crimes rates: namely motivation and opportunity. Our estimation results have suggested that in some cases, these two effects are equally strong. Therefore, without distinguishing between them, the coefficients of certain variables, such as unemployment rate, will remain ambiguous. Third, the empirical analyses in this chapter are based on the framework where crime rate is assumed to be affected only by its local factors. As offenders are mobile, the crime level in one area, however, could be affected by the conditions of neighbouring areas. Ignoring such “interaction” between neighbouring areas may potentially bias the effects of concerned variables. In order to incorporate such dependence between neighbourhoods, our next chapter will focus on detecting the spill-over effects of crime rates as well as relevant explanatory variables.

Chapter Five: Spatial Analyses

5.1 INTRODUCTION

Using a set of panel data, last chapter has provided a detailed analysis on the relations between various crime rates and other explanatory variables both in local areas over time, and across 43 police force areas in England and Wales. The assumption taken for such analysis is that each crime rate is only affected by local explanatory factors, and not influenced by any aspects from neighbouring areas. One may argue that this cannot comprehensively explain the complex relations between crime rates and factors such as unemployment, as crime rates in one area may be affected by crime rates and/or explanatory factors of its surrounding areas. In this chapter, we continue our analysis of factors affecting crime by looking at whether crime policies in one region affect crime in neighbouring regions. In particular, it tries to analyse whether crime rate in an area can depend not only on its local opportunities, but also on the opportunities in neighbouring areas. The rationale for hypothesizing such spatial spillover of crime is that, given people are mobile, an individual may choose to commit crime in a neighbouring area if the net benefit of doing so is higher than that from his home area. This could happen if, for example, the home area has tougher law enforcement (captured by a higher probability of detection) than that in the neighbouring area. In this case, some offenders will leave the home area due to fewer opportunities and commit crimes in neighbouring areas instead. This is the idea of crime displacement where the home area's policies displace crimes to neighbouring regions.[39]

[39] The following taken from the abstract of Measuring Criminal Spillovers: Evidence from Three Strikes (Tabarrock and Helland, 2009) summarises why economists regard this as a market failure.

This chapter firstly develops a simple theoretical model to analyse why people commit crimes and whether they do so in different regions. It then applies empirical analysis, by using panel data for 43 police force areas in England and Wales over the period 1998-2001, to estimate the strength and direction of such spillovers.

The model of crime spillover is constructed between two neighbouring areas to show how people allocate time between work and criminal activity. Spillovers are captured by allowing individuals to choose between home and neighbouring area to commit crimes. Committing crime in the neighbouring area may involve a higher cost which can be captured by a travel cost. Nevertheless, it could be more beneficial to do so if the opportunities are more abundant and/or the law enforcement is less tough in the neighbouring area. Tightening up the law enforcement in only one area may have the unintentional consequence of increasing the crime rate in the neighbouring area. Similarly, worse economic conditions in home area may also drive offenders to the neighbouring area for better opportunities. Therefore, this model will analyse the effects of variables such as detection rate and unemployment on the crime rates of both home and neighbouring area. Furthermore, this model will also predict the direction of these spill-over effects. For instant, the tightened law enforcement in one area will, in short-run, increase the crime rate in the neighbouring area through a negative spill-over effect. As time goes by, however, the crime rate of the neighbouring area could be reduced through the incapacitation effect. This is because tightened law enforcement leads to more crime prone people being caught and locked

---

California's Attorney General was pleased to announce that "An unintended but positive consequence of 'Three-Strikes' has been the impact on parolees leaving the state… The growth in the number of parolees leaving California is staggering." Law enforcement officer in other states were presumably less pleased. A displaced criminal is a benefit to California but a cost to other states. If such criminal spillovers are important, law enforcement will over-invest in policies that encourage displacement.

up and results in lower crime rates in both areas. As a result, the net effect will depend on the relative strength of the two effects.

After establishing this theoretical framework, the chapter will then apply spatial analysis to empirically test the inter-dependent relationship between the crime rates of contiguous areas. Specifically, spatial error and spatial lag models will be applied to study the spill-over effect of six types of crimes[40] with panel data covering 43 police force areas in England and Wales for the period 1998-2001. In addition, this chapter will also specify an empirical model to test whether crime rate is not only affected by local factors such as law enforcement and labour market opportunities, but also influenced by these characteristics of neighbouring areas.

The chapter is constructed as follows: section two reviews relevant literatures that try to explain the non-random spatial distribution of crime rates in different areas. Following the theoretical framework proposed in Fabrikant (1979), a simple model of crime spillovers is constructed to analyses one's decision on where to commit crimes: home or neighbouring areas. Hypotheses are then derived for later empirical test. Spatial analysis models are also included in this section which will be followed by their applications. Section three provides data description for both dependent and independent variables while section four tests the spatial distributions of these variables. Section five describes the empirical model under estimation and relevant statistical issues. This section also presents the estimated results and the implied interpretations. Section six summarizes the main results and the implications. It also

[40] The six types of crimes are violence against the person, sexual offences, robbery, burglary, theft and handling, and fraud and forgery.

discusses potential short-comings of this work on which may be improved in future researches.

5.2 LITERATURE REVIEW

5.2.1 Theoretical Perspectives

The uneven but non-random distribution of crime rate over space has been observed for a long time. Social scientists have developed different theories and techniques to explain such phenomenon. Anselin *et al.* (2000) has given a relatively comprehensive review on relevant theories methodologies. As summarized in their article, early sociological theories investigating the relationship between place and crime can be traced back to the middle of the 19th century: Guerry (1833, *cited in* Anselin *et al.* 2000) and Quetelet (1833, 1842, *cited in* Anselin *et al.* 2000), when they attempted to explain the differences in community crime levels with the social conditions of the resident population. Later on, such research made remarkable progress in the early 20th century thanks to the enormous contribution of the Chicago School. They obtained the record of each juvenile offender with their age, sex and home address and plotted on a map of Chicago. Based on the distribution of these juvenile offenses, Shaw and Mckay (1942) found a negative relationship between juvenile crime and the distance from central business districts.

More recent researches on crime distribution over space have greatly benefited from the invention of the Geographic Information Systems (GIS), which have made computerized mapping and spatial statistics possible. Adopting this technique, Curry and Spergel (1988) has found that delinquency is strongly correlated with poverty, while gang homicide, however, is predicted by the ethnic-race composition of local community based on the data from the U.S. Cohen *et al.* (1998) has also utilized the GIS system and investigated the distribution of gangs. They found that gangs are normally concentrated in neighbourhoods dominated by ethnic minorities. In addition, the "underclass", such as living in poverty and being unemployed, have also been positively correlated with gang activities.

Large proportion of papers analysing the distribution of crime rate have all, by some extend, benefited from two broadly cited theories, namely the routine activity theory and the social disorganization theory. The routine activity theory was initially developed in Cohen and Felson (1979) and later refined in Felson (1986, 1994). Brantingham and Brantingham (1993) have extended this theory to explain the distribution of crime. Instead of emphasizing the characteristics of offenders, this theory focuses on the circumstances in which they commit crimes. It argues that each successful offense requires the convergence in both time and space dimensions of three minimal elements: 1) an offender with both the intention to violate the law and the ability to carry out such action; 2) a person or object providing a suitable target; and 3) the absence of guardians capable to prevent such violation. Accordingly, the probability that an offense will occur at any specific time and place should be a function of the convergence of the three necessary elements. Any social conditions that affect the convergence of the three basic elements could possibly explain the variations in crime rate.

Place is essential in this theory in two ways. First, the physical features of a place can reduce the supervision effect that pedestrians could have. Newman (1972) has offered an example of this type. Public housings can increase population density in a building which could provide guardians against illegal activities, as the probabilities of people watching each others' back is higher in high-intensity estates. At the same time, however, because people are living vertically in public housings, such distribution can actually reduce the monitoring effect on each floor and weaken the informal social control among neighbours. Second, criminal activities are more likely to take place in target-rich environment. For example, thefts in a shopping mall, auto thefts from a large car park, or robberies in concentrated commercial areas (e.g. Engstad 1975; Brantingham and Brantingham 1982). In addition, certain activities such as alcohol consumption seem to be positively correlated with violent crimes (Roncek and Bell 1981); and abandoned buildings could attract illegal drug dealers. Based on the routine activity theory, therefore, it is not surprising that crime rate is not randomly distributed over space and certain crime types of crime tend to concentrate on certain places, namely crime hot spots.

The social disorganization theory is another one trying to explain the relationship between crimes and their occurring locations. It was developed in Shaw and McKay (1942), as they were attempting to establish a positive relationship between delinquency and the communities unable to conform to common values and to solve problems for the residents. The paper argues that delinquency is not a unique response of unique individuals; it is normal reaction by normal individuals to abnormal conditions. If a community cannot provide adequate protection for its residents and their properties and has to depend on outside agencies, some individuals will take the opportunity to conduct illegal activities at their will. In their empirical analysis, Shaw and McKay tried to explain the community-level delinquency rate with community-level economic conditions, ethnic heterogeneity and population turnover. Such explanatory factors are incorporated in hope to capture the instability and insecurity of communal environment. Their analysis proposes a spatial distribution pattern that juvenile delinquency rate is highest in inner-city areas and is decreasing with the distance away from city centre. As the analysis was carried out for the period of 1900-1933, large number of immigrants entered United States during that time and urban areas were the only places they could afford to live. Such fact implies that higher degrees of residential instability, ethnic diversity and social-economic deprivation are positively associated with the degree of urbanization and, hence, positively correlated with delinquency rate. Additionally, within inner-city areas, the probability of becoming an offender is associated with one's interpersonal network involving his family, gangs, and the neighbourhood. Finally, the degree of social and economic deprivation, population turnover and ethnic heterogeneity are all associated with social disorganization and hence, with crime.

The social disorganization theory has been followed by many latter papers. The degree of social disorganization is normally represented by five factors, namely, demographic, economic, social, family disruption and urbanization. Each of these factors can be measured by specific variables. For examples, economic status can be measured by various income levels; demographic conditions can be measured by the ethnic composition; family disruption can be measured by the percentage of single parent family; urbanization can be measured by population density and so on. Harries (1995) has found that poverty provides the strongest explanatory power for crime. Cahill and Mulligan (2003) has used ethnic composition, education, population density and other variables to measure the degree of social disorganization and tested their effects on violent crime. The results are generally consistent with expectation that higher degree of social disorganization is positively correlated with crime rate. One exception is that the population density does not exhibit significant effect as expected.

One common feature shared by the routine activity theory and the social disorganization theory is that they both attempt to explain the spatial variations in crime rates. More specifically, they seek to answer why the crime rates of certain areas are persistently higher than other areas? Instead of studying the behaviour of offenders, both theories focus on the characteristics of crime-prone locations. The aforementioned assumption by panel data analysis is that the crime rates of different areas are affected by specific features of those areas, but are independent from the features of neighbouring areas. Whereas, according to arguments presented above, crime rates could have spill-over effect across neighbouring areas. The crime rate of one area should, therefore, be affected by not only local relevant factors, but also such factors of neighbouring areas as well as neighbouring crime rates.

Fabrikant (1979) has developed a theoretical model which aims to derive an optimal allocation of police manpower when taking into consideration the possible spill-over effect of crime rate between neighbouring communities. This paper was motivated by two opposite opinions. On the one hand, Gylys (1974) suggests that “the residents of each political area have a positive marginal rate of substitution between another area’s consumption of police services and the goods that it consumes itself”. In other words, an increase in the police manpower in one area will not only reduce the local crime level, but also benefit the neighbouring areas. On the other hand, Press (1970) and Mehay (1977) oppose Gylys’s argument by providing empirical analyses suggesting that the increased law enforcement personnel in one area motivate criminals to spill over into adjacent areas. Having considered the opinions from both sides, Fabrikant (1979) tries to establish a theoretical model that is able to explicitly derive the effect of increased police manpower in one area on the crime rates of neighbouring areas.

The theoretical framework is constructed by incorporating the criminal spill-over effect into the theory of rational choice as developed in Becker (1968) and Ehrlich (1973). Potential criminals are assumed to be economically rational and attempt to maximize their expected utilities subject to their constraints. Thus, by allowing for people to commit crimes in both "home" and neighbouring areas, potential criminals will evaluate the expected punishment (subject to the probabilities of detection, conviction and imprisonment) and potential gain of illegal activities from not only the "home" area, but also evaluate these factors for neighbouring areas. Additionally, they need to evaluate the cost of travel in committing crimes in neighbouring areas. Potential criminals will be motivated to commit crimes in neighbouring areas by the expected gain over what they can get in their own areas, net the travelling cost and relative risk of committing crimes in neighbouring areas. The criminal spill-over equation between communities *i* and *j* can be derived by solving a system of supply-of-offenses and demand-for-control equations and expressed by the following function.[41]

$$O_{ij} = f(C_{ij}, W_{ij}, CR_{ij}, PCF_{ij}) \tag{5.1}$$

The dependent variable O_{ij} is the aggregated number of offenses committed in community *j* by offenders residing in community *i*. The independent variable C_{ij} represents the costs of committing crimes in community *j* when the

[41] For a complete derivation of the flows model see Fabrikant (Chs. 1, 2).

offenders are residing in community *i*. Such costs include the travel expenses and the time spent on travelling and carrying out the offenses. As it is difficult to exactly measure C_{ij}, it can be approximated by the distance the offender has to travel. W_{ij} represents the ratio between the potential gain from committing crimes in community *j* and that from committing crimes in community *i*. This variable measures how attractive committing crimes in community *j* is when the offenders are living in community *i*. CR_{ij} is defined as the ratio between CR_j and CR_i which represent the clear-up rates in community *j* and *i* respectively. For someone living in community *i*, higher CR_{ij} ratio represents relatively greater risk of getting detected in community *j* which will be less attractive as a result. The last independent variable PCF_{ij} is represented by the ratio between the number of potential offenses in community *j* and that in community *i* and measuring the relative competition pressure between communities *j* and *i*. The rationale is that higher number of potential offenses in a given community implies more fierce competition between offenders and thus lower marginal gain for an additional offense. When the competition pressure is higher in community *j* relative to that in community *i*, committing offenses in community *j* becomes less attractive with other variables being equal.

Given the definitions of the independent variables, their associations with aggregated offenses committed in community *j* by offenders from community *i* can be derived accordingly.

$$\frac{\partial O_{ij}}{\partial C_{ij}} < 0, \frac{\partial O_{ij}}{\partial W_{ij}} > 0, \frac{\partial O_{ij}}{\partial CR_{ij}} < 0 \text{ and } \frac{\partial O_{ij}}{\partial PCF_{ij}} < 0.$$

The above relationships suggest that the number of offenses spilling over to a neighbouring area depends on the relative conditions of both areas. Specifically, the spillovers of offenses into a neighbouring area is negatively correlated with the travelling expenditure (usually measured by travelling distance), positively correlated with the relative potential gain from neighbouring area, negatively correlated with the relative risk of detection of neighbouring area, and negatively correlated with the relative competition pressure (measured by the number of potential offenses) of neighbouring area.

5.2.2 A Model of Crime with Neighbourhood Effects

In this section, we construct a simple model of crime which allows the offenders to spillover into neighbouring areas. Unlike Fabrikant, we derive the criminal's choice explicitly from optimising behaviour. For simplicity, this model is developed upon a two-region situation: region 1 (home) and region 2 (neighbouring). Assume that there are in total N individuals living in region 1 and only a fraction of λ will commit crime under certain circumstances. The rest of the population, $(1-\lambda)N$, will never commit crime either because they are incorruptible or because they do not have the access to do so. People likely to commit crime, λN, will be generically labelled by i and individual i is allowed to choose between his home region (1) and the neighbouring region (2) to commit crime. Otherwise, he can choose to stay crime free. The trade-off between region 1 and 2 is that committing crime in the neighbouring region (2) will induce higher travel cost than remaining in home region (1). However, region 2 may have better opportunities for crime or lower probability of detection. For the moment, this model assumes that crime opportunities in each region are exogenously given as are the detection rates and punishment. The wage rate from legal sector is normalised to zero. Furthermore, the crime opportunities and wage rate are measured by how many units of a single consumption good they can buy. The punishment and travel cost are also measured by the units of the consumption good.

Let x and x^* denote the crime opportunities of region 1 and region 2 respectively in a single period; p and p^* denote the probability of detection in region 1 and region 2 respectively; θ denotes the punishment which is identical to region 1 and 2; and t denotes the travel cost of committing crime in region 2.

Given the notations, the expected return of criminal activity in region 1 is $x - p\theta$; and the expected return of criminal activity in region 2 is $x^* - p^*\theta - t$. In order to obtain non-zero crime level in region 1, it must satisfy $x - p\theta \geq 0$ given that the wage rate of legal sector has been normalized to 0. In addition, in order to make sure some corruptible people from region 1 will commit crime in region 2, it must satisfy $x^* \geq x$.

Given the above assumptions, a corruptible individual will commit crime in his home region (region 1) if

$$x - p\theta \geq x^* - p^*\theta - t\ .$$

Otherwise, he will travel to commit crime in the neighbouring region (region 2). Thus a critical value for the travel cost t can be derived which is

$$t_c = (x^* - x) + \theta(p - p^*)\ .$$

When $t \geq t_c$, the corruptible individual commits crime in the home region (region 1). Otherwise, he travels to the neighbouring region (region 2) to commit crime. If the travel cost t follows a distribution with cumulative distribution function F, then the crime rate in the neighbouring region (region 2) is given by $\Pr\{t \leq t_c\}$. By substituting the critical value of t, the crime rate of region 2 as a result of criminals migrating from region 1 becomes

$$u^* = F((x^* - x) + \theta(p - p^*))\ , \qquad (5.2)$$

which implies the crime in region 2 due to the criminal spillovers from region 1 is a function of parameters in both region 1 and region 2. In addition, the crime in region 1 is simply

$$u = 1 - F((x^* - x) + \theta(p - p^*)) . \tag{5.3}$$

As shown in the crime rate function of region 2, the crimes in regions 2 due to the criminal spillovers from region 1 is a function of relative criminal opportunities and detection rates between region 2 and region 1. Furthermore, this simplified theoretical model can be easily extended to incorporate punishment and labour market conditions varying across regions. For example, if the punishment θ is allowed to vary across region 1 and region 2, it will have similar effect as the detection rate. When the punishment in region 2 becomes relatively milder comparing to that of region 1, potential criminals could spillover into region 2 as a rational response.

Based on our model of crime spillovers, predictions can be derived for later empirical tests. As potential offenders prefer to commit crimes in opportunity-rich environment, regions with higher crime rates are likely to motivate offenders to spillover into neighbouring regions, because crime rate itself can reflect the criminal opportunities. When the criminal opportunities are exogenously given, higher crime rate in a region implies more fierce competition between potential offenders and some of them will spillover into neighbouring regions where the crime rates are lower. Thus, the crime rate of one region is expected to be positively correlated with the crime rates of its neighbouring regions. Such prediction will be tested in later part with spatial analysis models which allow the crime rate of one region depends on the crime rates of its neighbouring regions in addition to a set of independent variables.

Besides, the relative criminal opportunities between neighbouring regions can also be affected by a set of crime-influential factors. For example, higher income in one region will relatively increase the opportunities for crimes, especially property crimes, in this region to that of neighbouring regions. This can be appealing for offenders from neighbouring regions to spill over, hence, leave their own regions with fewer crimes. Likewise, higher detection rate in one region will relatively reduce criminal opportunities in this region to that of neighbouring regions. Thus, potential offenders may be more likely to spill into neighbouring regions due to better opportunities. Again, for the same reason, higher unemployment rate in one region could imply financial motivations for more offenders, hence creating fiercer competition between them. With more opportunities in neighbouring regions, some of them may relocate to neighbouring regions and contribute to crime rates of respective regions. These predictions can be tested through specifying an empirical model where the crime rate of one region is affected not only by its local explanatory factors such as detection rate, income level, unemployment rate and so on, but also by such factors from neighbouring regions. This model will also be tested with the data of England and Wales later.

It is worth noting that the analysis can be extended to non economic crimes (such as sexual offenses) by putting the act of committing these crimes in the utility function. However, the incentives still matter in that one would like to commit crime where it is relatively cheaper. Hence, the relative abundance of opportunities still matter for deciding where to commit crime as does relative costs.

The above model however captures the deterrence effect. What about the incapacitation effect? It might be argued that the incapacitation effect leading to higher penalties in one region leads to lowered crime in *both* regions. The intuition is as follows. If certain people are more crime prone then increased vigilance leading to more such people being locked up, hence, reduces the number of such crime prone people in the population and holding everything else constant reduces crime rate. However, one has to note that as a result of lower crime in region 1 (because of the increased penalty) fewer people are being caught (even though the fraction of criminals in region 1 who are caught increases). Therefore, the net impact of the incapacitation effect depends on whether the actual number of imprisoned people increases or decreases because of the increased vigilance. If indeed the incapacitation effect leads to more criminal types being locked up and the strength of that is bigger than the incapacitation effect, we would see a long run decline in crime in both areas as a result of increased vigilance in one area. Thus, one way to reconcile the opposing viewpoints in may be that the Press (1970) and Mehay (1977) were considering the deterrence effect while Gylys (1974) believed that the incapacitation effect would be the dominant effect.

5.2.3 Methodologies

There are two basic econometric models analysing the spatial dependence between neighbouring areas: the Spatial Autoregressive model (SAR) and the Spatial Moving Average model (SMA). Anselin (1988) has provided comprehensive discussion on both models. Given that the structure of these spatial analysis models and their relevant estimation strategies are quite standard, this chapter will draw extensively from Anselin (1988) when introducing the basic models. The main objective of this chapter is to apply the spatial analysis models to identify the spill-over effect of crime.

A general form of spatial dependence model is shown in the following equation.

$$
\begin{aligned}
& y = \rho W_1 y + X\beta + \mu \\
& \mu = \lambda W_2 \mu + \varepsilon \\
& \varepsilon \sim N(0, \sigma^2 I_n)
\end{aligned}
\tag{5.4}
$$

In equation (5.4), y represents an *nx1* vector of cross-sectional dependent variable and X is an *nxk* matrix of independent variables. W_1 and W_2 are known *nxn* spatial weight matrices that may take different forms, such as first-order contiguity relations, second-order contiguity relations, and functions of distances between areas etc. The first-order contiguity spatial weight matrix is the broadly used one which reflects the first-order contiguous relationships between different areas. Specifically, the elements of the first-order contiguity spatial weight matrix, w_{ij}, represents whether or not the two corresponding areas, *i* and *j*, are sharing common border (or point). While its main diagonal elements all have the value of zero, the off-main diagonal elements, w_{ij}, take either the value of one or zero. The value of one indicates that areas *i* and *j* are first-order neighbours (sharing common border) while the value of zero indicates otherwise. Practically, the spatial weight matrix is usually standardized so that for each *i*, $\sum_j w_{ij} = 1$. Such practice is for making sure that the spatial lag of, for example, dependent variable *Wy* equals to the average value of the dependent variables of neighbouring areas to a given area *i*. However, standardizing the spatial weight matrix is only optional and the non-standardised spatial weight matrix can also do the job.

By imposing restrictions on the model given by equation (5.4), two broadly used spatial dependent models can be derived. They are going to be discussed separately along with their estimation issues. By setting $W_2 = 0$, equation (5.4) can be simplified to the form:

$$
\begin{aligned}
y &= \rho W y + X\beta + \varepsilon \\
\varepsilon &\sim N(0, \sigma^2 I_n)
\end{aligned} \tag{5.5}
$$

This model is referred as the SAR model which is also called the spatial lag model for simplicity. Wy is the spatial lag of dependent variable y which gives the spatially-weighted average value of y in neighbouring areas. The inclusion of the spatial lag variable into the right hand-side of equation (5.5) can allow for spatial dependence existing in dependent variable y. In other words, the value of dependent variable y is determined not only by a set of explanatory variables, but also by the values of y from neighbouring areas. The coefficient of Wy, ρ, is measuring the strength of spatial dependence. In order words, ρ measures how much in the variation of dependent variable y can be predicted by the average value of y in neighbourhood areas.

However, the inclusion of the spatial lag of y will cause estimation problems for the Ordinary Least Squares (OLS) methodology, in that the estimated coefficients will be biased and inconsistent for both ρ and β. This is because the spatial lag term Wy is equivalent to an endogenous variable due to its correlation with the error term ε, even if ε is i.i.d. This fact violates one of the classic assumptions that independent variables should be uncorrelated with the error term for OLS method to generate unbiased and consistent estimates. Anselin (1988) has suggested a maximum likelihood approach which is able to provide consistent estimates. [42]Alternatively, the

[42] This maximum likelihood approach assumes the error term is normally distributed and the estimation can be carried out by a few steps: firstly, perform the OLS estimation for the model $y = X\beta_0 + \varepsilon_0$, which regresses the dependent variable on the independent variables from its own area; secondly, perform OLS estimation for the model $Wy = X\beta_L + \varepsilon_L$, which regresses the spatial lag of dependent variables on the same set of independent variables; thirdly, compute residuals from each of the previous regressions by $e_0 = y - X\hat{\beta}_0$ and $e_L = Wy - X\hat{\beta}_L$; fourthly, given e_0 and e_L find ρ that maximizes the concentrated likelihood function: $L_C = C - (n/2)\ln[(e_0 - \rho e_L)'(e_0 - \rho e_L)/n] + \ln|I - \rho W|$; And finally, given $\hat{\rho}$ that maximizes L_C, compute $\hat{\beta} = (\hat{\beta}_0 - \rho\hat{\beta}_L)$ and $\hat{\sigma}_\varepsilon = (1/n)(e_0 - \rho e_L)'(e_0 - \rho e_L)$.

spatial lag model can also be estimated by the Two-Stage Least Squares approach (TSLS) as suggested in Anselin (1988).[43] The choices of instruments for Wy follow the expectation of y conditional on X:

$$E[y \mid X] = (I - \rho W)^{-1} X\beta = X\beta + \rho WX\beta + \rho^2 W^2 X\beta + \ldots \tag{5.6}$$

As a result, the instruments will include a set of explanatory variables X as well as their spatial lags WX. In this way, Two-Stage Least Squares estimation will not be unbiased by the inclusion of Wy.

On the other hand, setting $W_1 = 0$ in equation (5.4) generates the SMA model, which is also known as the spatial error model, given by equation (5.7).

$$\begin{aligned} y &= X\beta + \mu \\ \mu &= \lambda W\mu + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{5.7}$$

The expression of SMA model shows that the error term μ is spatially autoregressive and the strength is measured by λ. This assumption could be reasonable if some influencing variables are unobservable or omitted from the estimation system and these omitted variables are spatially dependent. Since there is no endogenous variable, such as the spatial lag of y, included in equation (5.7) and the spatial dependence is only occurring to the error term, the OLS method will provide unbiased and consistent estimates for the desired coefficients β. However, the OLS estimation will be inefficient due the violation to one of the classic assumptions. For the OLS method to generate BLUE estimates, one assumption is that the error terms should be independent from one another with zero mean and constant variances. Conversely, equation (5.7) has shown that the error terms are spatially correlated and their

[43] The Two-Stage Lease Squares estimation has also been discussed in Anselin (1980, 1990), Land and Deane (1992), Kelejian and Robinson (1993), Kelejian and Prucha (1998).

variance/covariance matrix has inconstant variances on the main diagonal and systematic nonzero elements off the main diagonal. This can be demonstrated by the expression of the variance/covariance matrix given $\mu = (I - \lambda W)^{-1}\varepsilon$.

$$V = E(\mu\mu') = \sigma^2 (I - \lambda W)^{-1}(I - \lambda W)^{-1} \tag{5.8}$$

According to the features of the variance/covariance matrix as shown in equation (5.8), the estimated variances of the coefficients will tend to be smaller than they actually are and the hypothesis tests will exaggerate the significance of the coefficients.

It is not difficult to show that if λ, the autoregressive coefficient of the error term, were known, the regression given by equation (5.7) can be estimated by OLS with spatially filtered variables $y - \lambda Wy$ and $X - \lambda WX$:

$$y - \lambda Wy = (X - \lambda WX)\beta + \varepsilon \ , \tag{5.9}$$

where Wy and WX are the spatial lags of dependent and independent variables and the error term, ε, follows the classic assumptions of the OLS model. However, the λ coefficient is practically unknown and has to be estimated along with the regression coefficients β . Therefore, Anselin (1988) and relevant literatures have designed alternative estimation approaches which are able to avoid the influence of spatially correlated error terms. One of such approaches is the maximum likelihood estimation. By assuming the error terms are normally distributed, a likelihood function can be derived which needs to be maximized.[44]

[44] In the likelihood function given by equation $L = \sum_i \ln(1 - \lambda\omega_i) - n/2\ln(2\pi) - n/2\ln(\sigma^2) - (y - X\beta)'(I - \lambda W)'(y - X\beta)/2\sigma^2$, ω_i are the eigenvalues of the spatial weight matrix W. Since β and σ^2 can both be expressed as functions of the autoregressive coefficient λ, the above likelihood function can be reduced to the concentrated form which only contains λ as the unknown parameter as shown in equation (5.10). The estimated value of

The SAR model and the SMA model can also be combined with a set of spatial lags of independent variables. The extended SAR and SMA models can be expressed respectively by the following equations.

$$
\begin{aligned}
&y = \rho W y + X\beta + WX\gamma + \varepsilon \\
&\varepsilon \sim N(0, \sigma^2 I_n)
\end{aligned}
\qquad (5.11)
$$

And

$$
\begin{aligned}
&y = X\beta + WX\gamma + \mu \\
&\mu = \lambda W \mu + \varepsilon \\
&\varepsilon \sim N(0, \sigma^2 I_n)
\end{aligned}
\qquad (5.12)
$$

The extended SAR model given by equation (5.11) has similar features as the standard SAR model given by equation (5.5). One significant change, however, is that a set of spatial lags of independent variables can no longer be used as instruments in the Two-Stage Least Squares estimation since they have already entered the equation as independent variables. Consequently, a set of valid instruments is necessary for applying the Two-Stage Least Squares estimation. The maximum likelihood estimation, on the other hand, can still obtain consistent estimates for both the extended SAR and SMA models.

5.2.4 Applications

Due to rather limited numerber of literatures applying either the SAR or the SMA model on crime rate analysis, only two articles have been found for each case.

λ can simply be found by maximizing the concentrated likelihood function within the acceptable interval $1/\omega_{\min}$ and $1/\omega_{\max}$. $L_c = -n/2 \ln[(e'e)/n] + \sum_i \ln(1-\lambda\omega_i)$ (5.10)

In the concentrated likelihood function given by equation (5.10), $e'e$ is the residual sum of squares from equation (5.9): regressing the spatially filtered dependent variable on the spatially filtered independent variable.

Martin (2002) applies the SAR model to analyse the pattern of residential burglary in the city of Detroit based on the social disorganization theory. The spatial unit of analysis is the census tract and three hundred and twenty of them have entered the sample as proxies for neighbourhoods in Detroit. The dependent variable is the average burglary rate over the years 1995-1997. The independent variables have been selected according to the social disorganization theory. Initially, 11 variables have been chosen to represent the factors that influence the degree of social disorganization. However, as some of the variables are highly correlated, a principle components factors analysis has been conducted looking for a small number of linear combinations of these variables. Four factors have been produced as a result, namely, concentrated poverty, social capital, age composition and residential stability.

An OLS estimation has been performed as the first step. The results suggest that age composition is the strongest predictor for residential burglary rate, as it is positively correlated with the proportion of young people. The percentage population living in poverty also has positive effect on burglary rate as expected. The residential stability, though, has surprisingly obtained positive coefficient. The positive effect of residential stability does not necessarily mean that stable and familiar neighbourhood cannot improve the guardian effect that neighbours have for each other. It is possibly because more affluent and stable areas could be more appealing to potential criminals.

For further investigation, the SAR model is applied to detect the potential spatial dependence in burglary rate. The results have shown that the overall fit of the model has been substantially improved by introducing the spatial lag of the dependent variable. It has demonstrated that approximately 60 percent of the variation in burglary rate can now be explained comparing to 46 percent in the OLS estimation. Nevertheless, including the spatial lag of the dependent variable has reduced the magnitude of each coefficient for the independent variables. In addition, the residential stability becomes no insignificant as a result of including the spatial lag of burglary rate.

The application of SMA model has been given an example in Andresen (2006). The aim of this article is to analyse the spatial dependence in crime rates of different types using the data of Vancouver in 1996. The dependent variables under investigation are the crime rates of automotive theft, break and entering and violent crime. After separately mapping the crime rates over space with Geographic Information System software (GIS), the distributions of different types of crime rates have all shown very uneven patterns which have been taken as the evidence for applying the SMA model on the analysis. The census tract has been chosen to be the spatial analysis unit and there are totally 87 of them in Vancouver.

Initially, 13 independent variables have been selected to represent the factors influencing the degree of social disorganization such as ethnic heterogeneity, economic status, population composition and so on. Due to the potentially high correlation between these independent variables, Andresen has applied the general-to-specific method to reduce the influence of such correlation: for each type of crime rate, the model begins with including all independent variables; then the variable with the most insignificant coefficient will be dropped and the equation will be re-estimated. The same process will be repeated until all the remaining variables have significant coefficients.

The main findings of this paper can be summarized by the following points. Firstly, the variations in all three types of crime have been largely explained by the independent variables: 53, 65 and 78 percent respectively for automotive theft, break and entering and violent crime. Secondly, unemployment rate, young people percentage and the standard deviation of average family income have all exhibited positive correlations with the three crime rates in question. In particular, the unemployment rate has the strongest effect on these three crime rates. Thirdly, the proportion of single-parent households has positive impact on break and entering but shows no significant impact on automotive theft and violent crime. Finally, the ethnic heterogeneity shows significant and negative relationship with automotive theft and violent crime but insignificant correlation with break and entering. The negative effect of ethnic heterogeneity is opposite to the expectation that increasing diversity in ethnic composition will increase crime rate since it is supposed to measure the communal stability. One possible explanation could be the definition of this variable: ethnic heterogeneity has been defined as the percentage of recent (1981-1996) immigrants within the total census tract population. This is a different way of defining ethnic composition contrary to traditionally measurement on the proportions of different ethnic groups. Since the immigrants in Vancouver mainly containing economic-class, entrepreneurial-class and investor class, the proportion of them is not expected to directly affect crime rates such as automotive theft.

Martin (2002) and Andresen (2006) both try to explain the observed uneven and non-random distribution of crime rates with the spatial dependence model, either the SAR or the SMA. The unmentioned assumption behind both papers is that the non-random distribution of crime rates is caused by the spatial dependent effect. A major difference between these papers is that Martin (2002) applies the SAR model as it assumes the spatial dependence exists in the dependent variable. In other words, the crime rate in one region is affected not only by local explanatory variables, but also by the crime rates of neighbouring regions. This assumption is true when crimes can actually spill over into neighbouring regions due to the change in relative conditions of neighbouring regions. On the other hand, Andresen (2006) attributes the non-random distribution of crime to the feature of the error term. It assumes that the error terms are spatially dependent rather than normally distributed. Such assumption is reasonable when there are independent variables omitted from the equation and the omitted variables are spatially dependent.

Cahill and Mulligan (2007) and Malczewski and Poetz (2005) have also observed the non-random distribution of crime rates. Both papers attempt to explain this pattern with spatial heterogeneity. With visualization mapping equipment, it is found that the dependent variables in both papers, violent crime in Portland and residential burglary in London (Ontario) respectively, are highly concentrated at city centre areas. This clustering pattern reduces with the distance from city centre. Instead of assuming crime rates are spatially dependent, the two papers employ models allowing the coefficients of independent variables to vary over space. In other words, the crime rates of different regions are allowed to response differently to their local predictors. In order to estimate the coefficients varying across regions, both papers have employed the Geographically Weighted Regression (GWR) and compared the results to those generated by the conventional OLS regression. One of the most important propositions for this is that the GWR regression has substantially improved the explanatory power of the independent variables and some of the coefficients do display significant variations across neighbourhoods.

5.2.5 Weakness of Existing Empirical Literature

One similarity between the cited literatures and this chapter is that they have all noticed the uneven but non-random distribution of crime rates over space. Moreover, they are all attempting to explain such pattern with economic theories. There are, nonetheless, some advantages of this chapter comparing to other empirical literatures discussed above, especially Martin (2002) and Andresen (2006). Firstly, this chapter breaks down the overall crime rate into six categories according to the definitions given by the Home Office, as it acknowledges that different types of crimes have different features and thus probably respond differently to the same influencing factors. In contrast, the introduced empirical papers select only certain types of crime and some even select aggregated crime. None of them have provided comprehensive and yet systematic analysis for each individual type of crime. Secondly, this chapter has a unique data structure that is spatial-temporal. The data covers 43 police force areas in England and Wales over the period 1998-2001. Not only does such data structure increase the sample size and thus the degree of freedom for the analysis, it also makes it possible to introduce both year-specific and area-specific dummy variables. By doing so, the risk of omitting relevant but unobservable variables can be largely reduced. Thirdly, this paper applies both the SAR and the SMA models to detect spatial dependence, differing to both Martin (2002) and Andresen (2006), in which only one model is adopted. Additionally, this chapter also examines an extended model incorporating both independent variables and their spatial lags. The aim is that, by testing and comparing models as widely as possible, one is able to make an informed judgment on which model is the "best" to reflect the crime generating process.

5.3 DATA DESCRIPTION

Appendix IV contains a map showing the distribution of overall crime rate across police force areas in England and Wales for the year 2000/2001. As seen in the map, the highest crime rate occurs to the metropolitan areas such as Greater London, West Midlands, Great Manchester, and West Yorkshire. The second highest crime rate concentrates in the centre of England and Wales, created by regions of East Midlands and West Midlands. In contrast, most of coastal areas record a lower crime rate.

The dependent variables in this chapter are the six types of crime rates. The independent variables are chosen according to the classic theories of crime analysis developed in Becker (1968) and Ehrlich (1973). These influential factors include crime control, social-economic as well as demographic factors which can be represented by a set of variables which will be introduced in details. The data covers 43 police force areas in England and Wales for the time period 1998-2001.

5.3.1 Dependent Variables

The dependent variables under analysis are six types of crime rates defined by the Home Office, namely, violence against the person, sexual offences, robbery, burglary, theft and handling, and fraud and forgery. The first three categories have been defined as violent crimes given that they would threaten the physical well-beings of victims. The last three categories are defined as property crimes, since they are mainly targeting at personal properties, thus threaten victims' economic well-beings of.

Although robbery mainly aims at properties, it usually involves in physical attacks, hence, it has been treated as violent crime.

As mentioned previously, the crime rates under analysis are aggregated by police force areas and there are in total 43 of them across England and Wales. The time period being covered in this study is between 1998 and 2001. Such data structure is a unique feature of this thesis. As it is a spatial-temporal panel, it is possible to include both region-specific and year-specific fixed effects to account for the variables which might otherwise be correlated with the explanatory variables. This procedure can reduce the risk of fixed effects being included in the error term, thus avoid the estimation bias caused by simultaneous problem.

This spatial-temporal data structure, however, requires special arrangement when constructing the spatial weight matrix. As this chapter employs the first-order contiguity matrix, areas with common borders (or common point) will be considered as first-order contiguous areas and assigned with the value of 1 for the corresponding elements in the spatial weight matrix. Whereas, neighbourhoods with no common border (or common point) will be given the value of 0 in the spatial weight matrix. For England and Wales, the contiguous relationships between police force areas can be found in the map in *Appendix IV* and a 43*43 dimension first-order contiguity matrix can be constructed accordingly. However, due to the spatial-temporal data structure of this paper, the spatial weight matrix should have a dimension of 172*172 given that there are 43 police force areas over 4 years (43*4=172). In the spatial weight matrix, the police force areas are firstly arranged in alphabetical order. The organized police force areas are then stacked year by year from 1998 to 2001. Such organization method applies to both horizontal and vertical dimension of the spatial weight matrix. Next, the 43*43 dimension first-order contiguity matrix, which reflects the contiguous relationship between police force areas, is only applied to the sub-matrices on the main diagonal of the 172*172 dimension spatial weight matrix. The off-main diagonal sub-matrices have all zero values.

The detailed crime rates by police force areas were documented in *Criminal Statistics*[45] before and including the year 2000. It then has been transferred to *Crime in England and Wales*[46] after 2000. It is worth noticing that, up to and including year 2000 the crime rates were measured by the numbers of offenses recorded by police per 100,000 population. Year 2001 and onwards, the crime rates have been measured by the numbers of offenses per 1000 population. This change has reduced the accuracy in crime rates which has been discussed in last section. The statistic indicators on dependent variables are given in the following table. The number of observations in the last row gives idea on whether there is any missing value in the dataset. Since the data covers 43 police force areas over the period 1998-2001, there should be 172 observations (43 areas multiplied by 4 years) for each type of crime rate if no value is missing.

Table 5-1

Statistics for crime rates

		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	11,854	1,120	73	110
Median	8,827	860	63	64
Maximum	158,500	9,300	544	757
Minimum	4,760	380	0	0
Std. Dev.	18,605	1,211	59	143
Observations	172	172	172	172
		Property crimes		

[45] The serial numbers of *Criminal Statistics* are Cm4649, Cm5001 and Cm5312.
[46] *Crime in England and Wales, 2001/02.*

	Burglary	Theft and handling	Fraud and forgery
Mean	1,737	5,580	892
Median	1,486	3,649	416
Maximum	9,123	101,400	23,100
Minimum	400	1,400	200
Std. Dev.	1,134	11,774	2,733
Observations	172	172	172

5.3.2 Independent Variables

The independent variables include the detection rate, unemployment rate, real average weekly earnings and young people proportion. According to Becker (1968) and Ehrlich (1973), either higher probability of getting punished or more severe of punishment would reduce the incentive to commit crimes. This chapter however, only includes the detection rate as the law enforcement instrument because the data of prison population is only national and exhibits no variation across police force areas.

The social-economic statuses are represented by the unemployment rate and real average weekly earnings. As unemployment rate and average earnings could respectively represent the risk and potential income from legal labour market, Ehrlich (1973) has demonstrated their effects on crime rates: higher unemployment rate would have ambiguous effect on crime rate while higher average earnings could discourage criminal involvement.

The demographic factor is represented by young people proportion which refers to the percentage of 15 to 24 years old out of the total population. The idea is that young people, especially those less educated or unemployed, have been commonly regarded as more likely to commit crimes than other age groups. Since young people usually have relatively lower income by average than other age groups, their opportunity cost is lower. In other words, there is a bigger chance for them that the expected return from criminal activities would exceed how much they can get from the legal labour market. For the same reason, those less educated or unemployed young people are even more likely to commit crimes due to their lower opportunity cost.

Another reason to suspect that crime rate is affected by young people proportion is that the punishment for youth criminals is usually more lenient than that for adults. Since a stigma of criminal record is expected to greatly reduce the probability of employment and therefore reduce the expected payoffs from the legal labour market. As a result, this may foster repeated offenders among young people. Furthermore, for those under the age of 18, their criminal records will be sealed by the age of 18 and therefore will not affect their labour market outcomes. For this reason, it is arguable that the under 18s have even less to lose when committing crimes.

5.3.2.1 Detection Rate

The detection rate is measured by the proportion of recorded offences that have been cleared up. More specifically, the cleared up offences are referring to those cases when the offenders have been identified and given caution, fined or charged by the police. The data is disaggregated by police force areas and covers the period 1998 to 2001. The data source is also the annual command paper *Criminal Statistics* with serial numbers provided before. The following table gives the statistical summary as well as the information about any missing value.

Table 5-2

Statistics for detection rates

		Violent crimes		
	Overall crimes	Violence against person	Sexual offences	Robbery
Mean	29	76	65	31
Median	27	78	64	29
Maximum	69	97	103	93
Minimum	14	26	31	10
Std. Dev.	9	13	17	13
Observations	172	172	172	172

	Property crimes		
	Burglary	Theft and Handling	Fraud and forgery
Mean	17	22	44
Median	14	21	44
Maximum	56	54	86
Minimum	7	9	9
Std. Dev.	8	7	16
Observations	172	172	172

5.3.2.2 Unemployment Rate

The unemployment rate is constructed by dividing the number of unemployment benefit claimants by the number of workforce. The original data is on local authority level and has been aggregated according to police force areas. The data used here is from the website of nomis—the official labour market statistics.[47]

5.3.2.3 Real Average Weekly Earnings

This variable is measured by the deflated average weekly earnings for all industries and the original data is on local authority level. The same way has been applied to aggregate the original data into the frame of police force areas. The data is documented in *Annual Survey of Hours and Earnings*, collected from the website of national statistics.[48]

5.3.2.4 Young People Proportion

This variable is constructed by dividing the number of population between 15 to 24 years old by the population of all age groups. The original data is also estimated on local authority level and has been manually aggregated into police force areas according to geographic boundaries. The data source is also from the website of National Statistics.

[47] https://www.nomisweb.co.uk/Default.asp
[48] http://www.statistics.gov.uk/hub/index.html

The following table summarizes the key statistical indicators for unemployment rate, real average weekly earnings and young people proportion.

Table 5-3

Statistics for other independent variables

	Unemployment	Real average weekly earnings	Young People
Mean	3.04	183.14	11.68
Median	2.77	196.56	11.53
Maximum	7.11	509.81	13.44
Minimum	0.61	91.93	9.46
Std. Dev.	1.27	70.77	0.92
Observations	172	172	172

5.4 DETECTING SPATIAL DEPENDENCE

There are tests designed to detect the spatial distribution pattern of a given variable. Applying such tests can explicitly demonstrate whether the under testing variable is randomly distributed rather than displaying clustering pattern in the observations over space. The use of such tests can be very flexible. Before applying regression analysis, the dependent and independent variables can be tested to suggest whether spatial dependence presents and therefore spatial regression models should be considered. The residuals from different regressions can also be tested. For the OLS regressions, spatial dependence in the residuals would suggest that there could be omitted or unobservable variables remained which are spatially dependent. Hence, the specification of estimation should be modified with caution. Similarly, such tests can also be applied on the residuals generated by spatial regression models such as the SAR and the SMA. Such performance can check whether the spatial dependence in either the dependent variable or the error term has been successfully removed by the spatial regression model.

Although there are quite a few tests detecting spatial dependence, Moran's I and Geary's c are among the best known ones and are not difficult to compute. The Moran's I statistic can be computed by the following equation:

$$I = N / S_0 \sum_i \sum_j w_{ij} \cdot (x_i - \mu) \cdot (x_j - \mu) / \sum_i (x_i - \mu)^2 \ , \tag{5.13}$$

where N is the number of observations; w_{ij} is the element in the spatial weight matrix corresponding to the observation pair *i and j*; x_i and x_j are the observations for location *i and j* with mean μ; S_0 is a constant and takes the form $S_0 = \sum_i \sum_j w_{ij}$. For a row-standardized spatial weight matrix, S_0 equals N since S_0 is the sum of all weights and each row of the row-standardized spatial weight matrix sums to 1. In this case, the I statistic can be reduced to the following form:

$$I = \sum_i \sum_j w_{ij} \cdot (x_i - \mu) \cdot (x_j - \mu) / \sum_i (x_i - \mu)^2 \ . \tag{5.14}$$

On the other hand, the Geary's c statistic can be computed as following:

$$c = (N - 1) / 2S_0 [\sum_i \sum_j w_{ij} (x_i - x_j)^2 / \sum_i (x_i - \mu)^2] \ , \tag{5.15}$$

with the same notation as Moran's I statistic.

The statistical inference may be made according to a standardized *z*-value rather than the computed I or c statistics. The *z*-value may be constructed by subtracting the theoretical mean from the computed I or c statistic and dividing the result by the theoretical standard deviation. One thing worth noting is that the theoretical mean and standard deviation would both vary depending on the specific assumption made for the observation distribution. A common approach is to assume that the variable under testing follows a normal distribution. The *z*-value, in this case, would follow a standard normal distribution using the proper theoretical mean and standard deviation.

The statistical inference can be made by simply compare the computed z-value with the critical values of standard normal distribution. Another commonly used approach is to assume that each observation is equally likely to occur at all locations. In other words, the observations and their spatial arrangement are assumed to be irrelevant. This is referred as the randomization assumption. The *z*-value of this case also follows a standard normal distribution and the statistical inference can be made the same way as the assumption of normal distribution. The last approach is similar to the randomization assumption. Each observation is regarded to be equally likely to occurr at any location. The difference is that the mean and standard deviation of I or c statistic are generated empirically. Practically, this is carried out by randomly reshuffling the observations over all locations and re-computing the I or c statistic for each new sample. This is referred as the permutation approach and the mean and standard deviation for I or c statistic are then simply the computed moments.

Before conducting any regression analysis, all variables in this chapter, both dependent and independent, have been tested by both Moran's I and Geary's c techniques in order to get an idea about the distribution of observations over space. The 172*172 dimension spatial weight matrix has been row standardized before being applied in the tests and the results are summarised in the following tables.

Table 5-4

Dependent variables: crime rates

Variables	Moran's *I* Test			Geary C Test		
	Normal	Randomization	Permutation	Normal	Randomization	Permutation
Overall	0.04 (0.05)	0.04 (0.05)	0.04** (0.02)	0.55*** (0.06)	0.55*** (0.11)	0.55*** (0.12)
Violence against the person	0.16*** (0.05)	0.16*** (0.05)	0.16*** (0.03)	0.48*** (0.06)	0.48*** (0.10)	0.48*** (0.11)
Sexual offences	0.11** (0.05)	0.11** (0.05)	0.11** (0.04)	0.58*** (0.06)	0.58*** (0.11)	0.58*** (0.11)

Robbery	0.35*** (0.05)	0.35*** (0.05)	0.35*** (0.06)	0.57*** (0.06)	0.57*** (0.07)	0.57*** (0.08)
Burglary	0.11** (0.05)	0.11** (0.05)	0.11*** (0.04)	0.59*** (0.06)	0.59*** (0.09)	0.59*** (0.09)
Theft and handling	0.03 (0.05)	0.03 (0.04)	0.03* (0.02)	0.57*** (0.06)	0.57*** (0.11)	0.57*** (0.12)
Fraud of forgery	0.05 (0.05)	0.05 (0.04)	0.05* (0.02)	0.55*** (0.06)	0.55*** (0.12)	0.55*** (0.13)

The reported values in the table are the I or c statistic with the corresponding standard deviation in the brackets. In general, the results have shown evidences that the spatial dependence exists in each type of crime and it would be necessary to apply spatial regression models. In particular, the crime rates of four types, namely, violence against the person, sexual offences, robbery and burglary, have generated constantly significant statistics by rejecting all three assumptions in both Moran’s I and Geary’s c tests. For the rest of three types, theft and handling and fraud and forgery, the overall test statistics cannot reject the assumptions of normal and random distribution in Moran’s I test. However, for the permutation assumption and the whole Geary’s c tests, the computed statistics have been able to reject the null hypothesis of no spatial dependence existing. According to the test results, it would be reasonable to pay attention to the spatial dependent feature displayed by the dependent variables and such feature deserves explicit control.

The same tests have also been applied on the independent variables.

Table 5-5
Independent variables: detection rates

Variables	Moran’s *I* Test			Geary C Test		
	Normal	Randomization	Permutation	Normal	Randomization	Permutation
Overall	0.34*** (0.05)	0.34*** (0.05)	0.34*** (0.05)	0.71*** (0.06)	0.71*** (0.07)	0.71*** (0.07)
Violence against the person	0.28*** (0.05)	0.28*** (0.05)	0.28*** (0.05)	0.84*** (0.06)	0.84*** (0.06)	0.84*** (0.06)

Sexual offences	0.41*** (0.05)	0.41*** (0.05)	0.41*** (0.05)	0.61*** (0.06)	0.61*** (0.05)	0.61*** (0.05)
Robbery	0.29*** (0.05)	0.29*** (0.05)	0.29*** (0.05)	0.74*** (0.06)	0.74*** (0.07)	0.74*** (0.07)
Burglary	0.48*** (0.05)	0.48*** (0.05)	0.48*** (0.06)	0.51*** (0.06)	0.51*** (0.07)	0.51*** (0.07)
Theft and handling	0.37*** (0.05)	0.37*** (0.05)	0.37*** (0.05)	0.69*** (0.06)	0.69*** (0.07)	0.69*** (0.07)
Fraud of forgery	0.34*** (0.05)	0.34*** (0.05)	0.34*** (0.05)	0.73*** (0.06)	0.73*** (0.06)	0.73*** (0.06)

Table 5-5 shows that each type of detection rate is spatially dependent according to the overall highly significant test statistics. Similar conclusion can be made for the rest three independent variables: unemployment rate, real average weekly earnings and young people proportion. The results are given in table 5-6.

Table 5-6

Other independent variables

Variables	Moran's *I* Test			Geary C Test		
	Normal	Randomization	Permutation	Normal	Randomization	Permutation
Unemploy-ment rate	0.27** (0.05)	0.27*** (0.05)	0.27*** (0.05)	0.64*** (0.06)	0.64*** (0.06)	0.64*** (0.06)
Real earnings	0.80** (0.05)	0.80*** (0.05)	0.80*** (0.05)	0.11*** (0.06)	0.11*** (0.06)	0.11*** (0.07)
Young people	-0.12** (0.05)	-0.12** (0.05)	-0.12*** (0.05)	1.11* (0.06)	1.11** (0.05)	1.11** (0.05)

Aside from detecting spatial dependence, the test statistics can also tell whether the dependence is positive or negative. Given the obtained test results, all variables, both dependent and independent, have exhibited positive dependence over space except the young people proportion. Actually, the young people proportion has shown significant negative spatial dependence between neighbouring locations. Such inferences can be made by comparing the computed I or c statistic to their theoretical mean. For Moran's I test, the I statistic would indicate positive spatial dependence if it is higher than the theoretical mean; an I statistic lower than the theoretical mean would indicate otherwise. For Geary's c test, the case is opposite. The c statistic higher than theoretical mean indicates negative dependence while lower than theoretical mean indicates otherwise.

5.5 EMPIRICAL MODELS AND RESULTS

The basic empirical model under analysis is given by the following equation:

$$\begin{aligned} \ln(crime)_{it} &= c + \beta_1 \ln(\det ection)_{i(t-1)} + \beta_2 \ln(young)_{it} + \beta_3 \ln(earnings)_{it} \\ &+ \beta_4 \ln(unemployment)_{it} + dummy_{year} + dummy_{area} + \varepsilon_{it} \end{aligned} \tag{5.16}$$

The analysis of each type of crime will start with a conventional OLS estimation following the equation (5.15). One year lagged detection rate enters the equation instead of the contemporary detection rate to avoid the simultaneous problem. Both year-specific and area-specific dummies are included to control for the fixed effects which otherwise could be correlated with the independent variables. The inclusion of these dummy variables is an important advance on existing research, as most of which only employ cross sectional data and thus cannot control for the unobservable fixed effects. The other independent variables have been explained in the data description section.

Since heteroscedasticity often exists in cross sectional data (cross sectional-temporal data in this case), robust OLS estimations will be carried out if heteroscedasticity has been detected. The robust OLS estimations will not affect the magnitudes of coefficients: they will only correct the standard deviations of the coefficients and therefore generate the correct significant inferences.

In order to detect the spatial spillovers, the SAR and the SMA models will then be applied and the estimated results will be reported along with the results of standard and robust OLS estimations. Such arrangement will make it easy to observe how the estimated coefficients vary, in the sense of magnitudes and significance, across different models. In addition, it will also be easy to pick up the “best model” for the type of crime being analysed by comparing the goodness-of-fit of different models. Both SAR and SMA models will be estimated with the Maximum Likelihood (ML) method following the approaches suggested in Anselin (1988).

It is worth stressing is that there is no ready-to-use robust ML estimation in the software package to tackle heteroskedasticity in either SAR or SMA model. Instead, one feasible solution is to implement the Groupwise Heteroskedasticity (GHET) technique to reduce the influence of heteroskedastic variances of the error terms. However, applying such technique requires some knowledge about the possible cause of the heteroskedastic error terms. For example, the size, population or population density of different areas could all be the potential causes of heteroskedasticity and the GHET technique is based on such knowledge. Before applying the GHET technique, a categorial variable must be specified according to the possible reason of heteroskedasticity. This categorial variable only contains integer values and will divide the observations into different groups. If the categorial variable indeed reflects heterogeneous characteristics between groups, the variances of the error terms should be constant within groups and varying across groups. Therefore, by specifying a categorial variable and applying the GHET technique, the influence of heteroskedasticity will be reduced if the estimated groupwise variances are indeed unequal according to a Log Likelihood Ratio test (LR). A detailed application of GHET technique and categorial variable can be found in latter analysis.

As for further investigation, the crime rate being analysed will also be estimated by an extended equation that includes the spatial lags of independent variables. The specification is given by equation (5.17).

$$\begin{aligned} \ln(crime)_{it} &= c + \beta_1 \ln(\det ection)_{i(t-1)} + \beta_2 \ln(young)_{it} + \beta_3 \ln(earnings)_{it} \\ &+ \beta_4 \ln(unemployment)_{it} + \gamma_1 w\_\ln(\det ection)_{i(t-1)} + \gamma_2 w\_\ln(young)_{it} + \\ &\gamma_3 w\_\ln(earnings)_{it} + \gamma_4 w\_\ln(unemployment)_{it} + dummy_{year} + dummy_{area} + \varepsilon_{it} \end{aligned} \tag{5.17}$$

In this extended model, the crime rate of each area is assumed to be predicted not only by the independent variables of its own area, but also by the independent variables of its neighbouring areas. By applying the spatial weight matrix, the average values of the independent variables from neighbouring areas enter the equation (5.17) to represent their spill-over effects on the crime rate being analysed. As usual, both year and area dummies are included to count for the year-specific and area-specific fixed effects. The model specified by equation (5.17) will firstly be estimated by the standard OLS estimation. If heteroskedasticity is detected, two robust OLS estimations will be applied to correct the standard deviations of the coefficients and thus the significant inferences.

All variables will be taken logarithm before estimations except for sexual offences and robbery. As there are 0s in the crime rates, these two types of crime and their controlling variables will be estimated on their levels. The spatial weight matrix applied in both SAR and SMA models will be the row-standardized 172*172 dimension spatial weight matrix.

5.5.1 Burglary

Table 5-7 gives the estimation results for burglary following equation (5.15). Both year and area dummies are included in all the estimations. However, given there are in total 3 year dummies and 42 area dummies, their estimated coefficients are not reported in the table to save space.

Table 5-7

Burglary

Variables	OLS	Robust OLS		Spatial Error	Spatial Lag
		White	Jacknife		
Constant	2.19 (1.91)	2.19 (1.91)	2.19 (2.36)	1.85 (1.57)	0.13 (1.67)
Detection(t-1)	-0.06* (0.03)	-0.06* (0.03)	-0.06 (0.04)	-0.06** (0.02)	-0.06** (0.03)
Young People	0.40 (0.32)	0.40 (0.33)	0.40 (0.42)	0.48* (0.26)	0.41 (0.26)
Unemployment	0.14 (0.10)	0.14 (0.13)	0.14 (0.16)	0.11 (0.10)	0.12 (0.09)
Real Earnings	0.95*** (0.36)	0.95** (0.40)	0.95** (0.48)	0.99*** (0.30)	0.96*** (0.29)
Lambda	-	-	-	0.32*** (0.0006)	-
Spatial Lag	-	-	-	-	0.29*** (0.09)
Log Likelihood	256.79	256.79	256.79	261.35	260.50
Normality-Prob	0.62	0.62	0.62	-	-
Heteroskedasticity BP-Prob	0.0007	-	-	0.00004	0.0001
Heteroskedasticity Spatial BP-Prob	-	-	-	0.00004	0.0001
Spatial Lag Dependence-Prob	0.34	-	-	0.98	0.006
Spatial Error Dependence-Prob	0.13	-	-	0.003	0.03
Observations	172	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In OLS estimation, both spatial lag and spatial error dependences are diagnosed by the robust Lagrange Multiplier (LM) test which does not require the normality of the error terms. In SMA model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model.

As seen in the lower part of table 5-7, there are two tests for heteroskedasticity. The first one is either the Lagrange Multiplier test developed by Breusch and Pagan (the

BP test) or its studentized version suggested by Koenker and Bassett (the KB test). Which one to apply depends on the outcome of normality test. If the error terms are non-normal (for a significant level of 1%), the BP test will automatically be dropped and the KB test will be used instead. The second test for heteroskedasticity is only applied in spatial error and spatial lag models. As the BP test ignores the spatial dependence in the model, the second test is the spatially adjusted BP statistic which is based on the same principle. As shown in the table, they usually generate very similar results.

The reported statistics for the heteroskedasticity tests represent the statistical probabilities with the null hypothesis being the absence of heteroskedasticity. Given the results of such tests, heteroskedasticity constantly exists in the standard OLS, spatial error and spatial lag models. As heteroskedasticity usually underestimates the standard deviations of the coefficients and thus exaggerates their significances, the estimation results cannot be interpreted without controlling for heteroskedasticity. Two robust OLS estimations are therefore applied to correct the influence of heteroskedasticity in the standard OLS estimation and their results are also reported in table 5-7.

For both spatial error and spatial lag models, GHET technique is applied to generate unbiased inferences. As a categorial variable must be specified in the GHET technique, this paper has chosen the northern dummy to do the job. This is the result of a number of experiments. Different categorial variables, such as metropolitan dummy, population density dummy with different criteria and population size dummy with different criteria, have been applied separately hoping to capture the heteroskedasticity. None of those categorial variables, however, have generated significantly different groupwise variances. In contrast, the northern dummy has succeeded. The idea is that, by looking at the data, the northern areas on average have much higher unemployment rates and lower income comparing to the middle and southern areas. The crime rates, on the other hand, also appear to be higher in northern areas on average. Such distinctions between northern areas and the rest of England and Wales have inspired the selection of the categorial variable, which has indeed generated significantly unequal groupwise variances in most cases.

Table 5-8 gives the estimations of both spatial error and spatial lag models after correction of heteroskedasticity with the GHET technique. The corrected results are presented along with the original results in order to show how the coefficients vary before and after controlling for heteroskedasticity. All the estimations include the year and area dummies. Their coefficients are not reported to save space.

Table 5-8

Burglary

Variables	Spatial Error	Spatial Error GHET	Spatial Lag	Spatial Lag GHET
Constant	1.85 (1.57)	1.87 (1.55)	0.13 (1.67)	0.16 (1.66)
Detection(t-1)	-0.06** (0.02)	-0.06** (0.02)	-0.06** (0.03)	-0.06** (0.03)
Young People	0.48* (0.26)	0.49* (0.26)	0.41 (0.26)	0.43* (0.26)
Unemployment	0.11 (0.10)	0.10 (0.10)	0.12 (0.09)	0.11 (0.09)
Real Earnings	0.99*** (0.30)	0.98*** (0.29)	0.96*** (0.29)	0.94*** (0.29)
Lambda	0.32*** (0.0006)	0.36*** (0.09)	-	-
Spatial Lag	-	-	0.29*** (0.09)	0.31*** (0.10)
Log Likelihood	261.35	265.41	260.50	261.96
Heteroskedasticity BP-Prob	0.00004	-	0.0001	-
Heteroskedasticity Spatial BP-Prob	0.00004	-	0.0001	-
Spatial Lag Dependence-Prob	0.98	-	0.006	-
Spatial Error Dependence-Prob	0.003	-	0.03	-
Groupwise	-	0.02	-	0.09

Heteroskedasticity-Prob				
Observations	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model. The groupwise heteroskedasticity is diagnosed by LR test with the null hypothesis being equal groupwise variances.

By and large, the coefficients are quite robust over specifications given table 5-7 and 5-8. The lagged detection rate has negative and significant coefficient in all estimations except only the second robust OLS estimation. Such result is consistent with the expectation that the lagged detection rate could negatively affect crime rate through two channels: the potential criminals could learn that the detection rate had gone up from past experience, through friends or media reports. Realising such situation could deter the potential offenders to commit crimes by increasing their expected punishment from criminal activities. The second channel is through incapacitation. With more criminals being detected and put in jail, it is impossible for them to commit crimes in current time period.

Real earnings has also exhibited stable coefficient across estimations: positive and highly significant in all cases. This is not surprising since higher incomes can provide more opportunities for property crimes. For burglary, the main targets are those portable and valuable goods such as laptops, mobile phones and jewelleries etc from private dwellings. With higher incomes, people usually consume more on such goods and therefore provide more opportunities for burglary.

The coefficient of unemployment rate is also very robust across estimations. It shows positive but insignificant in all cases. Nonetheless, this result does not necessarily mean that unemployment rate has no significant correlation with burglary rate, as the reason of which could be complicated. The unemployment rate can have positive effect on crime rate, especially property crimes, since higher unemployment rate would reduce the expected return from legal labour market. The drop in the opportunity cost of committing crimes could be an incentive for some people to get involved in criminal activities. Given the nature of burglary, the crime takes place mostly in private dwellings. When unemployment rate is higher, more people would be staying at or near home in turn presenting fewer opportunities to potential criminals. Consequently, the unemployment rate could pick up both motivation and deterrence effects on burglary and give insignificant coefficient.

The coefficient of young people proportion is positive but insignificant in the OLS estimations. When spatial analysis models have been applied, young people proportion becomes positive and significant in most cases. The constantly positive but less significant effect of young people proportion could be explained by a combination of different reasons. On the one hand, more young people could increase the criminal supply given their lower opportunity cost to commit crimes than any other age groups. On the other hand, young people are also the targets of such crimes. Young people count the majority of college and university students most of whom possess laptops and mobile phones. They usually share houses with 2 or 3 housemates. Such houses are more attractive to potential criminals. This is because, firstly, the student houses are often located in less affluent areas and may not have anti-burglar measures installed. This makes them easier targets for burglars. Secondly, it is perhaps more profitable to attack student residences given that each student would possess at least one laptop and a mobile phone, as mentioned in previous chapter. Attacking one student house could obtain 2 to 3 laptops and mobile phones at one time, which may be more than twice the profit of attacking a normal residence. Given the above reasons, young people proportion is expected to have positive effect on burglary. However, burglary is typically a property crime and the fact is that young people have less to spend and possess less valuable goods than adults given their much lower incomes (if any). Although laptops and mobile phones are important targets of burglary, it is also true that potential burglars may be interested in other valuable goods that are often unaffordable by young people, such as high-end jewellery. Therefore, this fact could offset the positive effect of young people have on burglary and generate insignificant coefficient.

Both spatial error and spatial lag models have detected positive and significant spatial dependence. After applying the GHET technique to tackle heteroskedasticity, the spatial dependence is still positive and highly significant in both spatial analysis models. It is perhaps worth noting that the spatial dependence detected in the spatial error model has very different implication from that detected by the spatial lag model. Distinctively, the positive spatial dependence generated by the spatial error model implies that some influential factors have been left in the error term and such factors are spatially dependent. The spatial lag model, on the other hand, implies a spill-over effect of crime. When the burglary rate is high in a certain area, potential burglars will choose neighbouring areas to commit such crimes for relatively target-rich environments.

The goodness-of-fit of estimation is measured by the log likelihood ratio because the traditional R squared is not applicable to the spatial regression models. By presenting the log likelihood ratios, the goodness-of-fit of both spatial error and spatial lag models can be properly measured. In addition, it is also possible to directly compare the goodness-of-fit across estimations so that picking up the “best” model. According to the results given in table 5-7and 5-8, the spatial error model with GHET technique may be regarded as the best one in explaining the variations in burglary because of its highest log likelihood ratio. It is worth noting that the coefficient of real earnings is raised from 0.95 in the OLS estimations to 0.98 in the GHET spatial error model. Comparing to the coefficient of 0.94 in the GHET spatial lag model, the variable of real earnings becomes more powerful to explain the variations in burglary when the spatial dependence is controlled by the spatial error model. This phenomenon could be supportive to the above argument that the spatial error model with GHET technique performs better than the GHET spatial lag model in controlling the spatial dependence for burglary. In other words, the excluded explanatory variables are having greater spatial effects than the spill-over of burglary rate.

The proportion of young people does not exhibit strong impact on burglary (its coefficient is significant only on 10% level), although it shows positive effect as predicted by theory. Another way to capture the effect of young people is to incorporate the youth unemployment rate as an explanatory variable, instead of the proportion of young people. As young people have lower opportunity cost of committing crimes than older age groups, increased unemployment rate for the youth will further reduce their opportunity cost of involving in illegal activities and, therefore, increase the number of motivated potential offenders. Hence, we expect the variable of youth unemployment rate would exhibit more significant effect on the crime rate of burglary.

As an extended investigation, burglary rate has also been analysed according to equation (5.17). The aim of this is to test the relationship between burglary of an area and the independent variables of neighbouring areas as suggested by the theoretical model. Both year and area dummies are included in all estimations. Their coefficients are not reported for space-saving purposes.

Table 5-9
Burglary

Variables	OLS	Robust OLS	
		White	Jacknife
Constant	6.04 (4.44)	6.04 (4.68)	6.04 (5.72)
Detection(t-1)	-0.04 (0.03)	-0.04 (0.04)	-0.04 (0.04)
Young People	0.38 (0.33)	0.38 (0.31)	0.38 (0.41)
Unemployment	0.05 (0.15)	0.05 (0.20)	0.05 (0.26)
Real Earnings	0.94** (0.37)	0.94** (0.40)	0.94* (0.50)
Spatial Lag of Detection(t-1)	0.05 (0.07)	0.05 (0.11)	0.05 (0.14)
Spatial Lag of Young People	-1.28 (0.84)	-1.28 (0.91)	-1.28 (1.14)
Spatial Lag of Unemployment	0.27 (0.23)	0.27 (0.22)	0.27 (0.27)
Spatial Lag of Real Earnings	-0.23 (0.76)	-0.23 (0.86)	-0.23 (1.05)
Log Likelihood	259.67	259.67	259.67
Normality-Prob	0.38	0.38	0.38
Heteroskedasticity BP-Prob	0.0002	-	-
Observations	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level.

Looking at the results in the table above, the first thing needs to report is that the standard OLS estimation has generated heteroskedastic error terms according to the highly significant BP test. When the robust OLS estimations are applied, the estimation results do not change significantly, as real earnings is still the only significant predictor for burglary.

Comparing the results generated by equation (5.16) and (5.17), all previously included independent variables have shown the same signs and very similar magnitudes in the extended model as before. The difference is that the coefficients of lagged detection rate and young people proportion are no longer significant. Meanwhile, the spatial lags of independent variables have attained overall insignificant coefficients. Their signs though, are consistent with the expected spill-over effects.

The spatial lag of the lagged detection rate has positive coefficient in the extended model. It suggests that burglary is positively correlated with the averaged detection rate of neighbouring areas. This is because the burglary rate in one area will rise as a result of tightened-up crime control in neighbouring areas. This idea has been predicted in the theoretical model as the displacement effect of law enforcement. In a short run, strengthened law enforcement in one area will motivate its potential offenders to spillover into neighbouring areas, therefore increase the crime rates in neighbouring areas.

The spatial lag of unemployment rate has positive coefficient confirming with the prediction. Unemployment rate could have the motivation effect, as higher unemployment rate in neighbouring areas could motivate more potential offenders. The burglary rate of one area will, therefore, increase as a result of the offenders spilled over from neighbouring areas. In addition, as higher unemployment could reduce the opportunities for burglary, higher unemployment rate in neighbours will also drive the potential criminals to spillover therefore increase the burglary rate.

The negative coefficient of the spatial lag of real earnings can also be explained by the spill-over theory. Higher income levels in neighbouring areas could imply more opportunities available for burglars. As potential burglars would be attracted to neighbouring areas, the local burglary rate will be reduced accordingly.

The spatial lag of young people has negative coefficients in the extended model. This finding could support one of the assumptions explained previously. As young people, especially students, are usually regarded as the prime targets of burglary, higher young people proportion in neighbouring areas could create a target-rich environment attracting more burglars. When more burglars travel to neighbouring areas, their local areas will be left with lower burglary rate.

5.5.2 Theft and Handling

The same analysis procedure has been applied on theft and handling. For the concentrated model following equation (5.16), standard OLS estimation will be followed by robust OLS estimations if heteroskedasticity occurs. Spatial error and spatial lag models will be applied to test the spill-over effect of crime. When heteroskedasticity is detected, the GHET technique will be used to correct the biases. The extended model given by equation (5.17) will be applied to test the spill-over effects of independent variables. As usual, it will follow the same technique to control for heterodskedasticity. The year and area dummies are included all the time. Their coefficients, again, are not reported here to save space.

Table 5-10

Theft and handling

Variables	OLS	Robust OLS		Spatial Error	Spatial Lag
		White	Jacknife		
Constant	4.48*** (1.60)	4.48*** (1.55)	4.48** (1.93)	4.03*** (1.34)	2.38 (1.48)
Detection(t-1)	-0.13*** (0.05)	-0.13*** (0.04)	-0.13** (0.05)	-0.13*** (0.04)	-0.13*** (0.04)
Young People	-0.58** (0.26)	-0.58 (0.58)	-0.58 (0.87)	-0.49** (0.22)	-0.53** (0.22)
Unemployment	0.17** (0.09)	0.17* (0.10)	0.17 (0.12)	0.17** (0.08)	0.18** (0.07)
Real Earnings	1.21*** (0.30)	1.21*** (0.34)	1.21*** (0.45)	1.26*** (0.25)	1.19*** (0.25)
Lambda	-	-	-	0.24**	-

				(0.10)	
Spatial Lag	-	-	-	-	0.25*** (0.09)
Log Likelihood	288.06	288.06	288.06	290.39	291.07
Normality-Prob	0.64	0.64	0.64	-	-
Heteroskedasticity BP-Prob	0.0003	-	-	0.00001	0.0006
Heteroskedasticity Spatial BP-Prob	-	-	-	0.00001	0.0006
Spatial Lag Dependence-Prob	0.28	-	-	0.20	0.01
Spatial Error Dependence-Prob	0.82	-	-	0.03	0.85
Observations	172	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In OLS estimation, both spatial lag and spatial error dependences are diagnosed by the robust Lagrange Multiplier (LM) test which does not require the normality of the error terms. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model.

The diagnostic tests given in the table above suggest that heteroskedasticity exists in the standard OLS, spatial error and spatial lag models. Robust OLS estimations is applied to correct the standard errors of the coefficients in the OLS estimation. The GHET technique is adopted for both spatial error and spatial lag models with the northern dummy being used as the categorial variable. Table 5-11 reports the results of spatial error and spatial lag estimations both before and after applying the GHET technique. Because, in this way, it is easy to observe any change in the coefficients due to the application of GHET technique.

Table 5-11

Theft and handling

Variables	Spatial Error	Spatial Error GHET	Spatial Lag	Spatial Lag GHET
Constant	4.03*** (1.34)	4.08*** (1.33)	2.38 (1.48)	2.31 (1.48)
Detection(t-1)	-0.13*** (0.04)	-0.13*** (0.04)	-0.13*** (0.04)	-0.13*** (0.04)
Young People	-0.49** (0.22)	-0.42* (0.22)	-0.53** (0.22)	-0.47** (0.22)
Unemployment	0.17** (0.08)	0.18** (0.08)	0.18** (0.07)	0.19** (0.07)
Real Earnings	1.26*** (0.25)	1.21*** (0.25)	1.19*** (0.25)	1.14*** (0.25)
Lambda	0.24** (0.10)	0.28*** (0.10)	-	-
Spatial Lag	-	-	0.25*** (0.09)	0.28*** (0.09)

Log Likelihood	290.39	293.29	291.07	292.01
Heteroskedasticity BP-Prob	0.00001	-	0.0006	-
Heteroskedasticity Spatial BP-Prob	0.00001	-	0.0006	-
Spatial Lag Dependence-Prob	0.20	-	0.01	-
Spatial Error Dependence-Prob	0.03	-	0.85	-
Groupwise Heteroskedasticity-Prob	-	0.05	-	0.17
Observations	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model. The groupwise heteroskedasticity is diagnosed by LR test with the null hypothesis being equal groupwise variances.

Considering the results in table 5-10 and 5-11, the coefficients of independent variables are very robust across different regression models and agreeing with expectations in most cases. The lagged detection rate has negative and significant coefficient in all estimations. The magnitude, meanwhile, has been at -0.13 constantly. The unemployment rate has positive and significant coefficient in all estimations expect for the second robust OLS estimation. The magnitude of the coefficient is also very stable at either 0.17 or 0.18. The positive effect with unemployment has confirmed the expectation that higher unemployment would encourage some people to involve in crimes, theft and handling in this case, because of the lower expected legal income. Real earnings has also shown very stable coefficient that is positive and highly significant across all estimations. This is because real earnings has picked up the opportunity effect as it did in the analysis of burglary. With an increase in average income, people will probably carry more cash and spend more on goods. Such situation provides extra incentives and opportunities for potential thieves.

Contrary to expectation, young people proportion has constantly maintained negative coefficient in all estimations. Especially, the coefficient of young people proportion is negative and significant in the conventional OLS estimation. It then changes to insignificant in both robust OLS estimations. Such change is due to the exaggerated significance of young people proportion by heteroskedasticity. In both spatial error and spatial lag modes, the coefficient of young people proportion is always negative and highly significant either before or after applying the GHET technique. The problem is, however, the northern dummy does not successfully explain the heteroskedasticity in the spatial lag model according to the probability of 0.17 in the LR test. Therefore, it can only be concluded with confidence that young people proportion has negative and significant effect on theft and handling in the spatial error model, while cautions should be taken in interpreting its coefficient in the OLS and spatial lag estimations.

The negative coefficient of young people proportion (regardless its significance), may be due to the aforementioned double-edged effect that young people may have on theft and handling. Whilst higher young people proportion may encourage more potential thieves, given their lower opportunity cost, more young people may also increase the number of potential victims.

By looking at the sub-categories of theft and handling, it is not difficult to notice that theft from vehicle, as the largest sub-category, takes up to one third of the total number of theft and handling. It is probably reasonable to assume that theft from vehicle usually takes young people as potential targets, given that most young people cannot afford expensive cars that have better security measures, but they are more into gargets like CD players, SatNavs, customised speakers. This combined with cars' poor security features, such goods are attractive targets for thieves. Moreover, it is a general consensus that young people tend to be less cautious than matured adults. They are perhaps more likely to leave their personal belongings unattended in cars, which also provides opportunities for ”smash and grabs”. From this perspective, more young people could increase theft and handling by expanding the number of victims.

The second and third largest sub-categories are “other theft and unauthorised taking” and “theft or unauthorised taking of motor vehicle”. The common feature of these sub-categories is that the main targets are the older age groups who have relatively stable incomes and possess more valuable items and more expensive cars. An increase in the young people proportion could reduce the opportunities of such offences and hence the total of theft and handling. Taking this into consideration, then, the net effect of higher young people proportion remains ambiguous as all the possible effects analysed here may offer conflicting interpretations. In addition, we need to bear in mind that young people may not always be more prone to commit crimes than adults given their lower opportunity cost of doing so. As a criminal record will jeopardize the future labour market outcomes of the youth, they will be deterred to commit offences when considering this.

The spatial dependence in the error terms has been detected by the spatial error model both before and after applying the GHET technique. The estimated spatial dependent coefficient does not change much over the two cases. This result implies that there may be variables omitted from the spatially dependent specification. According to the spatial lag model, however, it cannot be concluded with confidence that theft and handling has spill-over effect. This is because, although both spatial lag estimations, with and without GHET technique, have attained positive and significant coefficient for the spatial lag of dependent variable, the northern dummy has not shown the power in explaining the heteroskedasticity. Hence, without solid evidence, it is reasonable to be sceptical that this type of crime would spill over.

The best performing model can be selected by comparing the log likelihood ratios across estimations, and the spatial error model with GHET technique has provided the highest log likelihood ratio and therefore the best fit for explaining the variations in theft and handling. A point worth the attention is that, when the spatial dependence is controlled by the spatial error model with GHET technique, the unemployment rate becomes more powerful in explaining the variations of theft and handing by exhibiting a coefficient of 0.18. (The coefficient of unemployment rate is 0.17 is OLS estimations). In contrary, the proportion of young people is less powerful in the GHET spatial error model: its coefficient is -0.58 in the OLS estimations and reduced to -0.42 when the spatial dependence in the error term is controlled for.

The estimation results of the extended model are given in the table below. Such model includes the spatial lags of independent variables as allowing them to have spill-over effects. As the BP test in the table below suggests the existence of heteroskedasticity, robust OLS estimations have been applied to generate reliable inferences. In order to save space, the coefficients of year and area dummies are not reported.

Table 5-12

Theft and handling

Variables	OLS	Robust OLS	
		White	Jacknife
Constant	12.25*** (3.63)	12.25*** (3.91)	12.25** (4.78)
Detection(t-1)	-0.09* (0.05)	-0.09* (0.05)	-0.09 (0.06)
Young People	-0.45 (0.47)	-0.45 (0.53)	-0.45 (0.82)
Unemployment	0.16 (0.12)	0.16 (0.14)	0.16 (0.17)
Real Earnings	1.16***	1.16***	1.16***

	(0.30)	(0.31)	(0.40)
Spatial Lag of Detection(t-1)	-0.07 (0.09)	-0.07 (0.09)	-0.07 (0.12)
Spatial Lag of Young People	-1.93*** (0.73)	-1.93** (0.80)	-1.93* (0.99)
Spatial Lag of Unemployment	0.11 (0.19)	0.11 (0.21)	0.11 (0.27)
Spatial Lag of Real Earnings	-0.70 (0.62)	-0.70 (0.68)	-0.70 (0.84)
Log Likelihood	295.09	295.09	295.09
Normality-Prob	0.69	0.69	0.69
Heteroskedasticity BP-Prob	0.0002	-	-
Observations	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level.

Comparing to the concentrated model, including the spatial lags of independent variables does not alter the signs of existing independent variables. The lagged detection rate and real earnings still have significant coefficients with the same signs. Young people proportion and unemployment rate, on the other hand, have gathered insignificant coefficients in the extended model.

The spatial lags of independent variables have broadly shown insignificant correlations with theft and handling with the spatial lag of young people proportion being the only exception. Particularly, the spatial lag of young people proportion has negative and significant effect on theft and handling. This result could be associated with the fact that young people, in addition to be more crime-prone, are also the main targets of certain crimes. As mentioned earlier, the sub-category of theft from vehicle takes up to one-third of the total numbers of theft and handling and is believed to mainly target on young people. As a result, with more young people in neighbouring areas, the potential criminals of theft from vehicle may well be drawn to the neighbouring target-rich environment and leave their own areas with lower crime rates.

Differing to the expectation, the spatial lag of lagged detection rate has obtained negative coefficient. This result advocates that the level of theft and handling will decrease as a result of tougher crime controls in neighbouring areas. As predicted in the theoretical model, the crime rate of one area should indeed have a negative relationship with its neighbouring crime control variables in long-run due to more criminals being locked-up. Consequently, the spatial lag of lagged detection rate could pick up the incapacitation effect, which may be beneficial in reducing crime rates for both local and neighbouring areas.

The spatial lags of both unemployment rate and real earnings have achieved the expected signs, positive for unemployment rate and negative for real earnings. As higher unemployment rate in neighbouring areas means more motivated thieves and fewer opportunities, theft and handling will increase as a result of thieves spill over from neighbouring areas. Likewise, higher income levels of neighbouring areas represent better opportunities for potential thieves. Theft and handling will fall as a result of criminal spillovers into neighbouring areas.

5.5.3 Fraud and Forgery

Table 5-13 and 5-14 present the estimation results for fraud and forgery following the concentrated specification given by equation (5.16). The year and area dummies are included in the all estimations as usual, but their coefficients are not reported here,

Table 5-13
Fraud and forgery

Variables	OLS	Robust OLS		Spatial Error	Spatial Lag
		White	Jacknife		
Constant	-0.96 (3.92)	-0.96 (4.58)	-0.96 (5.63)	-1.80 (3.31)	-2.12 (3.31)
Detection(t-1)	-0.13* (0.07)	-0.13* (0.07)	-0.13 (0.09)	-0.10* (0.05)	-0.11** (0.05)
Young People	-0.41 (0.65)	-0.41 (0.98)	-0.41 (1.43)	-0.45 (0.55)	-0.43 (0.55)
Unemployment	0.41* (0.21)	0.41* (0.24)	0.41 (0.29)	0.41** (0.19)	0.41** (0.18)

Real Earnings	1.80** (0.74)	1.80* (0.86)	1.80* (1.07)	1.97*** (0.62)	1.85*** (0.62)
Lambda	-	-	-	0.15 (0.10)	-
Spatial Lag	-	-	-	-	0.14 (0.10)
Log Likelihood	133.33	133.33	133.33	134.04	134.12
Normality-Prob	0.75	0.75	0.75	-	-
Heteroskedasticity BP-Prob	0.02	-	-	0.01	0.009
Heteroskedasticity Spatial BP-Prob	-	-	-	0.01	0.009
Spatial Lag Dependence-Prob	0.42	-	-	0.42	0.21
Spatial Error Dependence-Prob	0.60	-	-	0.23	0.79
Observations	172	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In OLS estimation, both spatial lag and spatial error dependences are diagnosed by the robust Lagrange Multiplier (LM) test which does not require the normality of the error terms. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model.

The diagnostic tests in the table above have shown that heteroskedasticity is a persistent problem across regressions. As robust OLS estimations have been applied to correct the standard OLS estimation, GHET technique is taken as the solution to deal with heteroskedasticity in both spatial error and spatial lag models. The results of GHET spatial error and spatial lag models are presented in the table below along with the original spatial models to show any change generated.

Table 5-14

Fraud and forgery

Variables	Spatial Error	Spatial Error GHET	Spatial Lag	Spatial Lag GHET
Constant	-1.80 (3.31)	-1.71 (3.29)	-2.12 (3.31)	-2.04 (3.29)
Detection(t-1)	-0.10* (0.05)	-0.11** (0.06)	-0.11** (0.05)	-0.12** (0.06)
Young People	-0.45 (0.55)	-0.35 (0.56)	-0.43 (0.55)	-0.33 (0.55)
Unemployment	0.41** (0.19)	0.44** (0.19)	0.41** (0.18)	0.43** (0.18)
Real Earnings	1.97*** (0.62)	1.91*** (0.62)	1.85*** (0.62)	1.81*** (0.61)
Lambda	0.15 (0.10)	0.13 (0.11)	-	-
Spatial Lag	-	-	0.14 (0.10)	0.13 (0.11)
Log Likelihood	134.04	136.03	134.12	135.10
Heteroskedasticity BP-Prob	0.01	-	0.009	-
Heteroskedasticity	0.01	-	0.009	-

Spatial BP-Prob				
Spatial Lag Dependence-Prob	0.42	-	0.21	-
Spatial Error Dependence-Prob	0.23	-	0.79	-
Groupwise Heteroskedasticity-Prob	-	0.14	-	0.16
Observations	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model. The groupwise heteroskedasticity is diagnosed by LR test with the null hypothesis being equal groupwise variances.

Adherent to the results shown in table 5-13 and 5-14, the variables that significantly affect fraud and forgery are the lagged detection rate, unemployment rate and real earnings. The lagged detection rate has negative and significant coefficient in all estimations except the second robust OLS regression. Therefore, it is rational to conclude that the problem of heteroskedasticity does not affect the significant correlation between lagged detection rate and fraud and forgery. As the lagged detection rate represents the lagged law enforcement effort, it can reduce fraud and forgery through both deterrence and incapacitation channels. Potential criminals could learn from their past experiences, their friends' experiences or media reports to estimate how likely they are going to get detected if commit such crimes. Therefore, higher detection rate of last time period would serve warnings on potential criminals by increasing their expected chance of punishment. For criminals who have been detected and sentenced into prison, it is impossible for them to recommit any crime while under custody.

The unemployment rate has overall positive and significant coefficient except again for the second robust OLS estimation. This result is in line with expectation since higher unemployment rate would reduce the expected return from legal labour market and hence the opportunity cost of criminal activities. Thus, higher unemployment may create incentives for engaging in property crimes such as fraud and forgery.

Real earnings has gained positive and significant coefficient in all estimations. This is because real earnings has picked up the opportunity effect as did it in previous analysis. People with higher incomes tend to spend more. Plus, with the increasing popularity of online shopping, those people may provide more opportunities for fraud and forgery. By checking the sub-categories of fraud and forgery, it is worth stressing that cheque and credit card fraud takes up to 90 percent of total offences in fraud and forgery. Therefore, it is not surprising to see that average income is positively correlated with fraud and forgery because of increased cheque and credit card payments from more affluent individuals.

One thing worth noting is that the GHET technique has failed to generate significantly different groupwise variances in either spatial error or spatial lag regression given the results of LR tests. In other words, fraud and forgery is not sensitive to whether or not the area is located in northern regions. In fact, this result does make sense due to the nature of fraud and forgery, as many of such activities are conducted remotely, which are not necessarily restricted by local unobservable characteristics. For example, credit card frauds are often carried out via the Internet, regardless geographical boundaries. In both spatial error and spatial lag models, no significant spatial dependence has been detected due to the presence of heteroskedasticity. As heteroskedasticity usually exaggerates the significance of coefficient, lacking of significance under heteroskedasticity is unlikely to become significant when heteroskedasticity is controlled. Hence, it is sensible to state that spatial dependence does not exist in either fraud and forgery or its omitted factors given that the Internet is the most commonly used platform and such crime is not as sensitive to geographic locations as other property crimes.

By comparing the log likelihood ratios, the spatial error model with GHET technique provides the best fit for fraud and forgery. Due to the problem of heteroskedasticity, however, the inferences of this model should be considered with scepticism.

The estimation results of the extended specification are given in table 5-15. The BP test proposes that heteroskedasticity exists in the standard OLS estimation. Thus, the robust OLS estimations are necessary to apply. All estimations have incorporated the year and area dummies while their coefficients are not reported.

Table 5-15

Fraud and forgery

Variables	OLS	Robust OLS	
		White	Jacknife
Constant	6.10 (8.75)	6.10 (8.28)	6.10 (10.45)
Detection(t-1)	-0.27*** (0.08)	-0.27*** (0.09)	-0.27** (0.11)
Young People	-0.76 (0.65)	-0.76 (1.13)	-0.76 (1.73)
Unemployment	0.49* (0.29)	0.49 (0.35)	0.49 (0.44)
Real Earnings	1.40* (0.74)	1.40* (0.83)	1.40 (1.05)
Spatial Lag of Detection(t-1)	-0.28*** (0.10)	-0.28** (0.11)	-0.28* (0.16)
Spatial Lag of Young People	1.67 (1.73)	1.67 (1.74)	1.67 (2.15)
Spatial Lag of Unemployment	-0.27 (0.45)	-0.27 (0.52)	-0.27 (0.65)
Spatial Lag of Real Earnings	-1.38 (1.51)	-1.38 (1.48)	-1.38 (1.80)
Log Likelihood	142.97	142.97	142.97
Normality-Prob	0.01	0.01	0.01
Heteroskedasticity BP-Prob	0.000002	-	-
Observations	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level.

While the previously included independent variables still have the same signs in the extended model, unemployment rate has become insignificant in both robust OLS estimations. Similarly, real earnings is significant only in the first robust OLS estimation and has lost its significance in the second robust OLS estimation. The lagged detection rate is the only one with significant coefficient among the previously included independent variables.

The spatial lags of the independent variables have broadly gathered insignificant coefficients. The spatial lag of lagged detection rate is the only one that has obtained significant coefficient with negative sign. Although it is possible that, in long-run, the crime rate in one area is negatively affected by the detection rates of neighbouring areas due to more criminals being locked-up, the interpretation of the negative relationship between fraud and forgery and the spatial lag of lagged detection rate needs more careful consideration. This is because, as argued previously, fraud and forgery is usually carried out via the platform of Internet and thus not as sensitive or restrictive to geographic locations as the other types of crime. Additionally, this point has been demonstrated by the previous analysis, in which either spatial error or spatial lag model has detected significant spatial dependence. Therefore, it would seem doubtful to explain the detected negative relationship between fraud and forgery and the lagged detection rate of neighbouring areas with the spill-over effect of tougher law enforcement, because the spill-over theory is based on the mobility of criminals. For the same reason, the coefficients of the other spatial lags of independent variables should not be explained by the spill-over theory either.

5.5.4 Robbery

Although robbery has been categorized as violent crime, the main purpose of such crime, however, is still to acquire valuable properties. As mentioned in the data description section, there are some 0s in the crime rate of robbery implying the number of robbery is less than 1 per 1000 population. To deal with this, the specifications given by equation (5.16) and equation (5.17) have been estimated based on the levels of both dependent and independent variables without taking logarithm. As usual, the year and area dummies are included all the time, with their coefficients not reported to save space.

Table 5-16
Robbery

Variables	OLS	Robust OLS		Spatial Error	Spatial Lag
		White	Jacknife		
Constant	-342.53 (229.70)	-342.53 (352.83)	-342.53 (453.44)	-349.75* (191.27)	-353.16* (190.64)
Detection(t-1)	-0.55 (0.69)	-0.55 (0.70)	-0.55 (0.88)	-0.61 (0.55)	-0.59 (0.57)
Young People	39.62** (16.72)	39.62 (24.71)	39.62 (32.08)	44.81*** (13.93)	42.74*** (13.88)
Unemployment	-25.31 (17.18)	-25.31 (27.45)	-25.31 (33.49)	-37.41** (14.53)	-28.89** (14.29)
Real Earnings	1.05*** (0.23)	1.05*** (0.34)	1.05** (0.52)	0.86*** (0.20)	0.88*** (0.20)
Lambda	-	-	-	0.25** (0.10)	-
Spatial Lag	-	-	-	-	0.20** (0.10)
Log Likelihood	-825.37	-825.37	-825.37	-822.79	-823.59
Normality-Prob	0.0000	0.0000	0.0000	-	-
Heteroskedasticity KB-Prob	0.00002	-	-	0.0000	0.0000
Heteroskedasticity Spatial BP-Prob	-	-	-	0.0000	0.0000
Spatial Lag Dependence-Prob	0.39	-	-	0.44	0.06
Spatial Error Dependence-Prob	0.17	-	-	0.02	0.06
Observations	172	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In OLS estimation, both spatial lag and spatial error dependences are diagnosed by the robust Lagrange Multiplier (LM) test which does not require the normality of the error terms. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model.

The heteroskedasticity tests given in the results table generate highly significant statistics indicating the existence of heteroskedasticity in the standard OLS, spatial error and spatial lag estimations. As usual, robust OLS estimations and the GHET technique have been applied to deal with this problem. While the results of robust

OLS estimations are given in table 5-16, the GHET spatial estimations are presented in the table below along with the normal spatial regressions.

Table 5-17
Robbery

Variables	Spatial Error	Spatial Error GHET	Spatial Lag	Spatial Lag GHET
Constant	-349.75* (191.27)	-351.31* (190.53)	-353.16* (190.64)	-353.66* (189.59)
Detection(t-1)	-0.61 (0.55)	-0.58 (0.54)	-0.59 (0.57)	-0.55 (0.56)
Young People	44.81*** (13.93)	42.68*** (13.87)	42.74*** (13.88)	40.94*** (13.78)
Unemployment	-37.41** (14.53)	-28.48** (13.87)	-28.89** (14.29)	-21.83 (13.68)
Real Earnings	0.86*** (0.20)	0.90*** (0.21)	0.88*** (0.20)	0.92*** (0.20)
Lambda	0.25** (0.10)	0.23** (0.10)	-	-
Spatial Lag	-	-	0.20** (0.10)	0.18* (0.10)
Log Likelihood	-822.79	-819.30	-823.59	-821.12
Heteroskedasticity BP-Prob	0.0000	-	0.0000	-
Heteroskedasticity Spatial BP-Prob	0.0000	-	0.0000	-
Spatial Lag Dependence-Prob	0.44	-	0.06	-
Spatial Error Dependence-Prob	0.02	-	0.06	-
Groupwise Heteroskedasticity-Prob	-	0.03	-	0.03
Observations	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model. The groupwise heteroskedasticity is diagnosed by LR test with the null hypothesis being equal groupwise variances.

The first thing needs to mention is that the GHET technique has indeed managed to control the heteroskedasticity on some level in both spatial error and spatial lag models. This is shown by the LR tests and the results of which have clearly rejected the null hypothesis of equal groupwise variances.

The estimated results in both tables presented above do not suggest significant correlation between robbery and lagged detection rate. However, the coefficient of lagged detection rate constantly has negative sign in all estimations. The other independent variables, meanwhile, have broadly shown significant effects on robbery. Real earnings, as usual, has positive and highly significant coefficient in all estimations. The explanation can be relatively straightforward. Real earnings has picked up the opportunity effect as higher incomes would generate more opportunities for robbery. This explanation is based on the assumption that people would consume more with higher incomes and such consumptions should include portable and valuable goods such as designer handbags, latest mobile phones and other digital gadgets.

The coefficient of unemployment rate is negative and insignificant in both normal and robust OLS estimations. When the spatial models are applied, the coefficient of unemployment turns negative and significant in both spatial error and spatial lag models. After applying the GHET technique, however, the coefficient of unemployment rate switches back to insignificant in the GHET spatial lag model whilst still remains significant in the GHET spatial error model. Regardless of its significance, the sign of this coefficient is constantly negative. Robbery is the type of crime that usually happens outdoors, and more often on the streets in less affluent areas. Moreover, timing is essential for successful actions. Determined robbers must come across with suitable victims at the right time and place. When unemployment rate rises, the probability will be reduced for potential robbers to meet suitable targets, as less people have to commute carrying with valuable items. This opportunity effect of higher unemployment rate will offset the motivation effect of unemployment rate and thus the net effect could be negative.

The significance of young people proportion is also varying across estimations. Basic OLS estimation has given positive and significant coefficient while the robust OLS estimations have corrected it back to insignificant. In the case of spatial analysis, the positive and significant coefficient of young people proportion is not affected by applying the GHET technique to both spatial error and spatial lag models. The sign of young people proportion, nonetheless, has been positive all the time. The constantly positive sign validates with the expectation that more young people could imply more potential criminals given their lower opportunity cost of committing crimes. At the same time, due to the double-edged effect, more young people could also increase the number of potential victims. As youngsters are more into chasing the latest trend in fashionable and digital products, they are perhaps popular targets in robbery. Furthermore, as discussed above, young people may not as cautious and caring to their personal possessions as their older counterparts, they may be again more exposed to the risk of being robbed than other age groups. By summarising the above reasons, young people should be positively correlated with robbery. Despite this, one should not forget the fact that age is negatively correlated with income level in most cases. Young people, on average, have lower incomes and possess less valuable goods. Furthermore, the potential future labour market payoff could make the youngsters less keen to become offenders. These effects could somehow offset the positive effect that young people proportion has on robbery.

According to both spatial error and spatial lag regressions with GHET technique, location really matters in the case of robbery, with positive and significant spatial dependence being detected. Nevertheless, the interpretation of which should depend on whether the spatial dependence presents in the error terms or in the dependent variable. The significant spatial dependence in the spatial error model connotes the omission of spatially dependent variables from the specification. The spatial lag model, on the other hand, suggests a spill-over effect that robbery has over neighbourhoods.

By comparing the log likelihood ratios, the spatial error model with GHET technique has provided the strongest explanation power for robbery.

The table below reports the results of the extended model incorporating the spatial lags of independent variables. As the standard OLS estimation has the problem of heteroskedasticity, robust OLS estimations are applied to correct the coefficient inferences.

Table 5-18

Robbery

Variables	OLS	Robust OLS	
		White	Jacknife
Constant	186.17 (468.51)	186.17 (507.96)	186.17 (671.98)
Detection(t-1)	-0.58 (0.69)	-0.58 (0.70)	-0.58 (1.01)
Young People	41.49** (16.73)	41.49 (28.82)	41.49 (42.99)
Unemployment	-35.95** (17.48)	-35.95 (27.22)	-35.95 (36.21)
Real Earnings	0.90*** (0.30)	0.90 (0.60)	0.90 (1.04)
Spatial Lag of Detection(t-1)	-0.07 (1.51)	-0.07 (1.29)	-0.07 (1.67)
Spatial Lag of Young People	-66.84* (37.28)	-66.84 (50.15)	-66.84 (70.14)
Spatial Lag of Unemployment	54.56* (31.47)	54.56 (50.93)	54.56 (78.31)
Spatial Lag of Real Earnings	0.66 (0.63)	0.66 (1.07)	0.66 (1.47)
Log Likelihood	-818.21	-818.21	-818.21
Normality-Prob	0.000001	0.000001	0.000001
Heteroskedasticity BP-Prob	0.00002	-	-
Observations	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level.

As shown in the above table, none of the independent variables have displayed significant coefficients in the robust OLS estimations. Whilst the previously included independent variables still have the same signs as in the concentrated model, they have all lost their significant correlations with robbery rate.

The spatial lags of independent variables are not able to provide solid evidence for the predicted spill-over effects of independent variables. Firstly, the spatial lag of lagged detection rate has negative and highly insignificant coefficient. This result suggests that the robbery rate in one area is highly unlikely to be affected by the lagged detection rate of neighbouring areas. Consequently, this result is not supporting the predicted spill-over effect of detection rate. Secondly, the spatial lag of young people proportion has negative but insignificant coefficient. The negative coefficient could imply that the robbery rate of one area would be reduced as there are more young people in neighbouring areas. This result will be valid only if, as argued previously, young people also represent the prime victims of robbery. Therefore, more young people in neighbouring areas may indicate more opportunities for robbery and so attract potential robbers to spillover from nearby areas. Thirdly, the spatial lag of unemployment rate has shown positive sign as expected although the coefficient is significant. This result implies a positive correlation between the robbery rate of one area and the unemployment rate in neighbouring areas. As higher unemployment could entail more motivated potential robbers as well as fewer opportunities for them, those determined robbers are likely to spillover into neighbouring areas for better opportunities and in doing so, push up the robbery rate of neighbouring areas. Finally, the spatial lag of real earnings has found positive sign in the extended model contrasting to prediction. Although higher incomes may denote more opportunities for robbery, higher income level in neighbouring areas is predicted to reduce the robbery rate of the area under study by drawing potential robbers away to more affluent neighbouring areas. As a result, the positive coefficient is not supporting the predicted spill-over effect of real earnings.

5.5.5 Sexual Offences

Sexual offences is a typical type of violent crime and seems have no direct relationship with social-economic factors. It is usually believed that people commit sexually offences owing to their difference in "personality" or "taste". However, as we mentioned earlier while the crime itself may generate utility, potential criminals may still try to commit the crime where it is easier to do so and the deterrence effects may apply here as well. The same analysis procedure has therefore been applied on sexual offences in hoping to generate interesting insights.

The crime rate of sexual offences and independent variables enter the estimations on their levels without taking logarithm. This is because there are 0s in the crime rate indicating there is less than 1 sexual offence per 1000 population. Both year and area dummies are included in each estimation. Their coefficients are not reported for space limit.

Table 5-19
Sexual offences

Variables	OLS	Robust OLS		Spatial Error	Spatial Lag
		White	Jacknife		
Constant	1136.1*** (196.80)	1136.1** (535.41)	1136.1 (762.58)	1070.52*** (161.54)	1127.9*** (164.34)
Detection(t-1)	0.10 (0.33)	0.10 (0.48)	0.10 (0.66)	0.16 (0.27)	0.09 (0.28)
Young People	-82.22*** (14.16)	-82.22* (44.38)	-82.22 (62.69)	-79.12*** (11.65)	-81.22*** (11.80)
Unemployment	-8.34 (15.13)	-8.34 (15.58)	-8.34 (19.64)	-1.98 (12.35)	-5.75 (12.61)
Real Earnings	-0.57*** (0.20)	-0.57 (0.76)	-0.57 (1.20)	-0.51*** (0.16)	-0.56*** (0.17)
Lambda	-	-	-	-0.17 (0.12)	-
Spatial Lag	-	-	-	-	-0.20* (0.11)

Log Likelihood	-799.63	-799.63	-799.63	-798.91	-798.50
Normality-Prob	0.0000	0.0000	0.0000	-	-
Heteroskedasticity KB-Prob	0.0000	-	-	0.0000	0.0000
Heteroskedasticity Spatial BP-Prob	-	-	-	0.0000	0.0000
Spatial Lag Dependence-Prob	0.35	-	-	0.29	0.13
Spatial Error Dependence-Prob	0.56	-	-	0.23	0.66
Observations	172	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In OLS estimation, both spatial lag and spatial error dependences are diagnosed by the robust Lagrange Multiplier (LM) test which does not require the normality of the error terms. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model.

For the diagnostic tests in the table above, heteroskedasticity presents in the normal OLS estimation as well as both spatial analysis models, so that heteroskedasticity correction is necessary before the results can be interpreted. The results of robust OLS estimations are given in the table above along with the normal OLS estimation. The results of GHET spatial error and spatial lag models are reported in the table below, which can be easily compared to the ordinary spatial models.

Table 5-20
Sexual offences

Variables	Spatial Error	Spatial Error GHET	Spatial Lag	Spatial Lag GHET
Constant	1070.52*** (161.54)	924.80*** (161.32)	1127.9*** (164.34)	989.38*** (164.59)
Detection(t-1)	0.16 (0.27)	0.10 (0.28)	0.09 (0.28)	0.01 (0.28)
Young People	-79.12*** (11.65)	-67.61*** (11.67)	-81.22*** (11.80)	-70.03*** (11.85)
Unemployment	-1.98 (12.35)	-1.91 (11.89)	-5.75 (12.61)	-6.05 (12.13)
Real Earnings	-0.51*** (0.16)	-0.41** (0.17)	-0.56*** (0.17)	-0.46*** (0.17)
Lambda	-0.17 (0.12)	-0.19* (0.11)	-	-
Spatial Lag	-	-	-0.20* (0.11)	-0.20* (0.12)
Log Likelihood	-798.91	-795.72	-798.50	-796.57
Heteroskedasticity BP-Prob	0.0000	-	0.0000	-
Heteroskedasticity Spatial BP-Prob	0.0000	-	0.0000	-
Spatial Lag Dependence-Prob	0.29	-	0.13	-
Spatial Error Dependence-Prob	0.23	-	0.66	-
Groupwise Heteroskedasticity-Prob	-	0.04	-	0.05

Observations	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model. The groupwise heteroskedasticity is diagnosed by LR test with the null hypothesis being equal groupwise variances.

The results have shown that the most influential predictors for sexual offences are young people proportion and real earnings. The coefficient of young people proportion is negative and highly significant in the normal OLS estimation as well as the spatial regressions, as still is the case in the first robust OLS estimation. The second robust OLS estimation has failed to provide significant coefficient. The negative relationship between young people proportion and sexual offences may entail that such crime is not very sensitive to young age from the perspectives of both potential offenders and victims. The incentives of committing such crimes are probably genetic and do not necessarily decrease as one gets older. There should be a clear distinction between the type of people who are naturally sexually offensive and others who may never accept such behaviours. This distinction is certainly not the age. Speaking of potential victims, females are obviously the majority, due to their physiological characteristics. Although one cannot deny that some sexual offenders are more target-specific, such as paedophiles, others are seemingly more of an opportunist, with the age of the victims not a determine factor. Perhaps, timing and location are more critical for those offenders.

Real earnings, as a typical social-economic factor, constantly exhibits negative correlation with sexual offences. The coefficient is significant in the normal OLS estimation and turns to insignificant in both robust OLS estimations. In the spatial analysis models, the coefficient of real earnings is negative and highly significant both with and without the GHET technique being applied. The negative effect of real earnings on sexual offenses could be explained that given higher incomes, the potential victims (mainly women) would be able to spend more on self-protections, such as taking cabs in late night or live in a less crime-prone area, and reduce the probability of contacting potential criminals.

The lagged detection rate and unemployment rate have insignificant coefficients in all estimations. Interestingly, the coefficient of unemployment rate is constantly negative. Considering sexual offences mainly committed outdoors, higher unemployment rate reduces the unavoidable movement of people on streets. Moreover, higher unemployment also has inevitably financial implications for the unemployed. As less money is around, they might go out at night as much as they did before. As a result, for opportunist sexual offenders, relatively fewer targets may be available.

Negative spatial dependence has been detected by both GHET spatial error and spatial lag regressions. In fact, the coefficient of spatial dependence in the spatial error model becomes significant when introducing the GHET technique. The northern dummy has indeed captured some of the heteroskedastic feature according to the LR tests. In the spatial lag model, the negative and significant spill-over effect of sexual offences is consistent with expectation. It is the “personality” or “taste” that distinguishes the potential criminals from others. When potential criminals relocate to neighbouring areas seeking for opportunities, their residential areas would be left with less people with sexually offensive intentions and hence less such incidents.

The log likelihood ratios have shown that the spatial error model with GHET technique has the best performance in explaining the variations in sexual offences.

The table below gives the estimation results of the extended model. As usual, both year and area dummies are included all the time.

Table 5-21
Sexual offences

Variables	OLS	Robust OLS	
		White	Jacknife
Constant	693.97* (390.33)	693.97 (680.83)	693.97 (1145.44)
Detection(t-1)	-0.03 (0.34)	-0.03 (0.36)	-0.03 (0.51)
Young People	-89.46*** (13.92)	-89.46* (51.37)	-89.46 (84.28)
Unemployment	-17.43 (15.06)	-17.43 (21.29)	-17.43 (29.80)
Real Earnings	-0.49* (0.25)	-0.49 (0.83)	-0.49 (1.49)
Spatial Lag of Detection(t-1)	1.31* (0.61)	1.31 (1.14)	1.31 (1.78)
Spatial Lag of Young People	12.10 (32.32)	12.10 (51.22)	12.10 (80.91)
Spatial Lag of Unemployment	90.25*** (27.26)	90.25 (85.29)	90.25 (142.51)
Spatial Lag of Real Earnings	0.66 (0.52)	0.66 (0.84)	0.66 (1.39)
Log Likelihood	-789.56	-789.56	-789.56
Normality-Prob	0.0000	0.0000	0.0000
Heteroskedasticity BP-Prob	0.0000	-	-
Observations	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level.

As robust OLS estimations are applied to control for heteroskedasticity, none of the independent variables have significant coefficients expect young people proportion in the first robust OLS estimation. As usual, including the spatial lags of independent variables does not alter the signs of previously included independent variables. Their significances, however, have been greatly reduced. While the coefficients of young people proportion and real earnings are constantly significant in the concentrated model, their significant effects on sexual offences have been eliminated in the extended model.

The spatial lags of independent variables have, once again, indicated overall insignificant coefficients. Their signs are consistent with expectation in most cases. Firstly, the spatial lag of lagged detection rate has gained positive sign in the extended model. This result confirms the prediction that, in short-run, sexual offences in one area will increase as a result of higher detection rate in neighbouring areas, because tougher law enforcement is expected to drive potential offenders into nearby areas. Secondly, the spatial lag of young people proportion has gained positive coefficient entailing a positive relationship between the sexual offences in one area and the young people proportion in neighbouring areas. On the other hand, as analysed previously, this type of crime is generally not sensitive to young age but should depend on the offender's personality instead. Higher young people proportion does not necessarily denote more potential offenders who might spillover into neighbouring areas. Thirdly, the spatial lag of unemployment rate is positively correlated with sexual offences agreeing with both previously findings and the predicted spillovers of potential offenders. In the concentrated model, the unemployment rate is negatively correlated with sexual offences albeit with an insignificant coefficient. This is because higher unemployment rate may provide fewer opportunities for sexual offences given that such crimes normally happen at public places. In a model where potential offenders are allowed to spillover into neighbouring areas, higher unemployment rate in neighbouring areas could drive the potential offenders to spillover for better opportunities and thus increase the level of sexual offences in the area under study. Finally, the spatial lag of real earnings has also shown positive sign consisting with previous finding. The negative correlation between real earnings and sexual offences in the concentrated model implies that higher incomes could provide extra protections for potential victims in some cases. In an area with higher income and less opportunities, potential offenders are likely to spillover into nearby areas and increase their levels of sexual offences.

5.5.6 Violence against the Person

The same procedure has been applied on violence against the person searching for less obvious relationships with social-economic factors. Estimations firstly follow the specification given by equation (5.16) and the results are reported in table 5-22. As usual, year and area dummies are included all the time to count for year-specific and area-specific fixed effects. Both dependent and independent variables have been taken logarithm before estimations because there is no 0 in the crime rate.

Table 5-22

Violence against the person

Variables	OLS	Robust OLS		Spatial Error	Spatial Lag
		White	Jacknife		
Constant	3.34 (3.60)	3.34 (4.63)	3.34 (5.74)	3.12 (3.04)	3.21 (3.08)
Detection(t-1)	-0.23* (0.12)	-0.23* (0.13)	-0.23 (0.17)	-0.24** (0.10)	-0.23** (0.10)
Young People	-0.009 (0.59)	-0.009 (0.59)	-0.009 (0.76)	0.03 (0.50)	-0.002 (0.50)
Unemployment	0.37* (0.20)	0.37* (0.21)	0.37 (0.26)	0.36** (0.17)	0.37** (0.14)
Real Earnings	0.87 (0.67)	0.87 (0.83)	0.87 (1.01)	0.90 (0.56)	0.87 (0.56)
Lambda	-	-	-	0.04 (0.11)	-
Spatial Lag	-	-	-	-	0.02 (0.11)
Log Likelihood	150.74	150.74	150.74	150.82	150.75
Normality-Prob	0.0000	0.0000	0.0000	-	-
Heteroskedasticity KB-Prob	0.005	-	-	0.0000	0.0000
Heteroskedasticity Spatial BP-Prob	-	-	-	0.0000	0.0000
Spatial Lag Dependence-Prob	0.34	-	-	0.30	0.85
Spatial Error Dependence-Prob	0.31	-	-	0.69	0.30
Observations	172	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In OLS estimation, both spatial lag and spatial error dependences are diagnosed by the robust Lagrange Multiplier (LM) test which does not require the normality of the error terms. In SEM model, the spatial lag dependence is based on the LM test and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model.

Based on the standard OLS estimation, it seems that only the lagged detection rate and unemployment rate have significant effects on violence against the person. Their coefficients are negative and positive, respectively. As robust OLS estimations are introduced to correct the impact of heteroskedasticity, the first robust OLS estimation still generates significant coefficients for lagged detection rate and unemployment rate with the same sign. The second robust OLS estimation, whereas, produces overall insignificant coefficients for all four independent variables. In the same way, the results of spatial error and spatial lag models are corrected for heteroskedasticity before interpretation. The results of GHET spatial error and spatial lag models are presented in table 5-23.

Table 5-23

Violence against the person

Variables	Spatial Error	Spatial Error GHET	Spatial Lag	Spatial Lag GHET
Constant	3.12 (3.04)	2.33 (2.85)	3.21 (3.08)	2.60 (2.90)
Detection(t-1)	-0.24** (0.10)	-0.29*** (0.11)	-0.23** (0.10)	-0.29*** (0.11)
Young People	0.03 (0.50)	0.44 (0.50)	-0.002 (0.50)	0.39 (0.50)
Unemployment	0.36** (0.17)	0.36** (0.17)	0.37** (0.14)	0.36** (0.17)
Real Earnings	0.90 (0.56)	0.91* (0.53)	0.87 (0.56)	0.86 (0.53)
Lambda	0.04 (0.11)	0.05 (0.11)	-	-
Spatial Lag	-	-	0.02 (0.11)	0.01 (0.11)
Log Likelihood	150.82	156.90	150.75	155.80
Heteroskedasticity BP-Prob	0.0000	-	0.0000	-
Heteroskedasticity Spatial BP-Prob	0.0000	-	0.0000	-
Spatial Lag Dependence-Prob	0.30	-	0.85	-
Spatial Error Dependence-Prob	0.69	-	0.30	-
Groupwise Heteroskedasticity-Prob	-	0.002	-	0.001
Observations	172	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level. In SEM model, the spatial lag dependence is based on the LM and the spatial error dependence is based on the LR test. In the SAR model, the spatial lag dependence is based on the LR test and the spatial error dependence is based on the LM test. Lambda in the first part of the table represents the spatial dependence coefficient of the error term in the spatial error model. Spatial lag in the first part of the table represents the included spatial lag of dependent variable in the spatial lag model. The groupwise heteroskedasticity is diagnosed by LR test with the null hypothesis being equal groupwise variances.

Table 5-23 shows that the northern dummy categorial variable has indeed explained the heteroskedasticity on some level in the results of LR tests. Such results advocate that, by dividing the police force areas into two groups according to the northern dummy, the groupwise variances of error terms are indeed unequal that is the cause (or at least partly the cause) of detected heteroskedasticity. By applying the GHET technique, the results of both spatial error and spatial lag models are now corrected for heteroskedasticity and more reliable to interpret.

The coefficient of lagged detection rate is constantly negative and significant except for the second robust OLS estimation. This result proposes that lagged detection rate can reduce violence against the person as expected. Nevertheless, this specification cannot separately identify the deterrence and incapacitation effect of detection rate given that the lagged detection rate may reflect both. Potential offenders can estimate their probability of detection based on past experiences and records. Thus, higher lagged detection rate will deter such crimes by increasing the offenders' expectation of detection rate. On the other hand, there is usually a time gap between the detection and conviction of real offenders. Higher lagged detection rate implies more offenders being detected in last period and being convicted and being currently put in prison. As a result, lagged detection rate may reduce violence against the person through incapacitation effect.

The coefficient of unemployment rate is positive and significant in all estimations except again for the second robust OLS estimation. This result implies that higher unemployment rate is associated with higher violent crimes again person. Although this result is consistent with the expectation that higher unemployment rate could increase crime level by reducing the expected legal payoff, the violent crimes are quite different in nature from property crimes. This is because the gain from violent crimes (excluding robbery) is difficult to calculate in monetary term, as they are not seeking for financial gains. Therefore, the positive effect of unemployment rate could be through less obvious channel rather than the simple monetary trade-off between legal and illegal activities. For example, unemployment is often associated with social and domestic disorders. The psychological impact of losing jobs combining with increased economic pressure may instigate violent conducts against persons.

The coefficient of real earnings shows up as significant only in the GHET spatial error model, but has been constantly positive under all circumstances. The positive sign of real earnings could be explained by the fact that more than 90 percent cases of violence against the person belong to the sub-category of possession of weapons. If weapons, such as guns, are regarded as consumption, it will be not difficult to understand the positive correlation between income levels and weapon possessions and hence, the total level of violence against the person.

As shown in both GHET spatial error and spatial lag estimations, no significant spatial dependence has been detected. This finding may be supported by the nature of such crime that violence against the person is quite spontaneous and the potential offenders are less likely to travel around looking for suitable targets. Moreover, one shall also acknowledge that, many cases of violence against the person are fuelled by inner city gang (or quasi-gang) rivalries. As these specific targets are between rival groups of gang, their violent activities are normally confined within one local area. By comparing the log likelihood ratios, the GHET spatial error model is the best one to explain the variations in violence against the person.

The results of the extended estimations are summarized in the table below. Without reporting the coefficients, both year and area dummies do enter all estimations.

Table 5-24
Violence against the person

Variables	OLS	Robust OLS	
		White	Jacknife
Constant	16.95**	16.95*	16.95
	(8.40)	(8.73)	(10.64)
Detection(t-1)	-0.20	-0.20	-0.20
	(0.13)	(0.14)	(0.18)
Young People	0.01	0.01	0.01
	(0.60)	(0.62)	(0.87)
Unemployment	0.26	0.26	0.26
	(0.27)	(0.28)	(0.34)
Real Earnings	1.16*	1.16	1.16
	(0.69)	(0.78)	(0.97)
Spatial Lag of	0.24	0.24	0.24
Detection(t-1)	(0.19)	(0.20)	(0.30)
Spatial Lag of	-4.24***	-4.24**	-4.24*
Young People	(1.61)	(1.75)	(2.17)
Spatial Lag of	0.22	0.22	0.22
Unemployment	(0.42)	(0.52)	(0.65)
Spatial Lag of	-1.35	-1.35	-1.35
Real Earnings	(1.44)	(1.48)	(1.79)
Log Likelihood	157.24	157.24	157.24
Normality-Prob	0.0000	0.0000	0.0000
Heteroskedasticity BP-Prob	0.01	-	-
Observations	172	172	172

*** represents significant at 1% level; ** represents 5% level; * represents 10% level.

In the extended model incorporating the spatial lags of independent variables, none of the predictors have obtained significant coefficients in the robust OLS estimation except for young people proportion. In fact, the coefficient of young people proportion is negative and highly significant in both robust OLS estimations. This result demonstrates that violence against the person is decreasing as the young people proportion in neighbouring areas increases. One possible explanation is that, if young people represent the prime targets of such crime, more young people in neighbouring areas create a target-rich environment. According to the spill-over theory of crime, such target-rich environment will attract potential criminals from neighbouring areas and thus leave their home areas with fewer crimes. However, this explanation is questionable due to the nature of violence against the person. As argued in above. It is therefore not convincible enough to explain the effect of spatial lag of young people proportion simply with the spillovers of potential criminals. The interpretation of this result deserves further analysis.

The lagged detection rate and unemployment rate become insignificant in the extended model, whereas they are significant in the concentrated model. Given the largely insignificant coefficients, the predicted spill-over effects of independent variables are not supported in the case of violence against the person.

5.6 CONCLUSION

The aim of this chapter is to test the spill-over effects of crime rates as well as a set of crime-influential variables. This work is motivated by the idea that, given the mobility of individuals, one is able to choose where to commit crime. Where exactly to commit crime depends on the expected returns and cost relative to other places.

This chapter has developed a simple model including two regions, namely, home and neighbouring. By allowing individuals to choose between the two regions for criminal activities, it has shown that the crime in the neighbouring area as a result of criminals spilling over from the home area is positively correlated with the relative criminal opportunities in the neighbouring area and negatively correlated with the relative crime control in the neighbouring area.

In order to test the predicted spill-over effects of both crime rates and crime-influential variables, this chapter has specified two empirical models. Model 1 (the concentrated model) assumes that the crime rate in one area is only predicted by its local crime-influential variables. However, the crime rate is allowed to depend on the crime rates of its neighbouring areas. The rationale of this model is that crime opportunities can be proxied to an extent by crime rates, higher crime rate in one area implies more fierce competition between criminals and thus potentially fewer opportunities for each of them. As criminals are mobile, neighbouring areas with lower crime rates will attract them to spillover due to the relatively better crime opportunities. As a result, the crime rates of neighbouring areas will increase as a result of the criminal spillovers.

The concentrated model has been tested with panel data covering 43 police force areas in England and Wales over the period 1998-2001. The predicted positive spill-over effect of crime rate has been strongly supported by the data employed. Particularly, in the case of burglary, theft and handling, robbery, and sexual offences, the crime rates have shown significant correlation with the crime rates of neighbouring areas. On the other hand, the crime rates of fraud and forgery and violence against the person do not exhibit spatial dependence between the crime rates of neighbouring areas. This finding is highly consistent with the varying natures across different types of crime. For burglary, theft and handling, robbery, and sexual offences, finding a suitable target at the right time and the right place is essential. Therefore, potential criminals will travel and choose locations for better crime opportunities. Fraud and forgery is not sensitive to locations because the majority of which are committed via the Internet. The main reason for violence against the person not being sensitive to location is its occurring is mostly spontaneous. In summary, as for the crimes where locations really matter, the predicted spill-over effect of crime rate has been strong supported by empirical analysis.

The empirical model 2 (the extended model) is designed to test the spill-over effects of the independent variables. In this model, the crime rate of one area is predicted by not only its local independent variables, but also the independent variables from neighbouring areas. The rationale is that the crime opportunities in one area can affected by its local crime control policies, labour market conditions and so on. Specifically, tightened-up crime control should reduce crime opportunities in one area. Likewise, better economic conditions, such as higher income level or lower unemployment rate, should provide better crime opportunities. In a model where criminals are mobile, neighbourhoods with more crime opportunities are supposed to appeal criminals from neighbouring areas, thus, affect the crime rates of neighbouring areas.

Testing model 2 using the same data set, however, has not generated supportive results. In fact, the crime rate in one area is not significantly correlated with the independent variables of neighbouring areas in most case. Furthermore, by incorporating the independent variables of neighbouring areas into the model, the effects of local independent variables have been greatly eliminated. In most cases, the independent variables with significant coefficients in model 1 have displayed insignificant coefficients in model 2.

Nonetheless, the overall insignificant results of model 2 should not be concluded that the crime rate in one area is not affected by the crime-influential factors in neighbouring areas. One possible explanation for these unsatisfactory results shown in model 2 may be that police force areas are still too big as proxies of neighbourhoods. This provides great incentives for further researches. In the future, it may be useful to test the spill-over effects of crime-influential factors with more disaggregated data. For example, the cross-sectional or panel data based on contiguous neighbourhoods within a city could be a good choice.

Chapter Six: Conclusion

6.1 Evaluation of Results

This thesis contains (apart from the literature review) three (mainly) empirical chapters. The first two chapters (i.e. chapters 3 and 4) examine the relationships between the empirically testable counterparts of Becker and Ehrlich's crime influential variables on different types of crime rates by applying it different data sets and using different econometric techniques. The third empirical chapter aims to test whether there is spatial dependence between the crime rates of neighbouring areas controlling for relevant explanatory variables. In this chapter, we are going to briefly summarise our empirical analyses in chapter 3-5 as well as our main findings. We will also discuss the limitations of each chapter and further research prospects.

In Chapter Three, we start our empirical analyses by carrying out simple time series analysis. The aim of this chapter is to examine the correlations between property crimes and unemployment as well as law enforcement from the angle of temporal variations. We construct the empirical model following the theoretical frameworks in Becker (1968) and Ehrlich (1973). The crime rates being analysed are overall crime rate, burglary, theft and handling, and fraud and forgery. Each of these crimes is predicted by unemployment, detection rate and custody rate with the last two variables being proxies for law enforcement instruments. According to both models in Becker (1968) and Ehrlich (1973), we expect both detection rate and custody rate to negatively affect crime rates due to their deterrence and incapacitation effects that they have as law enforcement instruments. The effect of unemployment on crime is ambiguous to predict as argued in both Ehrlich (1973) and Cantor and Land (1985) and shown in a large number of empirical papers.

Our time series analyses are performed by applying cointegration tests and error correction mechanism (ECM) on national level data covering 1971-2000. Due to the influence of new counting rules on crime rates introduced in 1998, we find that both overall crime rate and fraud and forgery are affected by this change. Therefore, in the analyses of these two crimes, we restrict the time period being examined to only include 1971-1997. Our first step leading to cointegration test is conducting unit root test for each variable and determining their order of integration. The results shows that all the variables in our analyses are integrated of order one indicating they are non-stationary on levels but stationary once differenced. Such results enable us to move on to cointegration tests between different crimes and their related variables. The results from cointegration tests suggest that, firstly, unemployment is positively cointegrated with overall crime, burglary, and theft and handling, while it is negatively cointegrated with fraud and forgery. These results indicate that, in the long-run, unemployment affects overall crime, burglary and theft through its motivation (i.e. by lowering employment opportunities it motivates criminal activity) effect. Such positive correlation could also be caused by the effect of crime on unemployment: higher crime rates imply more participation in illegal activities of offenders and probably less time spent on legal labour market; furthermore, criminals could have worse labour market opportunities due to either criminal records or deficient skills. The positive cointegration between unemployment and fraud and forgery seems to match one's expectation for this type of crime. This is because people usually need jobs to obtain opportunities for such crimes. Secondly, custody rate has negative cointegration with all the crime rates being examined, suggesting its deterrence and incapacitation effect as law enforcement instrument. And thirdly, the crime-specific detection rate is negatively correlated with overall crime rate which is consistent with prediction. On the other hand, detection rate has positive cointegration with burglary, theft and handling, and fraud and forgery. This result can be explained by the fact that higher crime rates require tougher crime controls.

Given the obtained cointegrations, we are able to estimate a short-term dynamic ECM model for each crime. We find that the change in custody rate has been the strongest predictor for the change in each crime rate with constantly negative effect. This result is supportive of both theoretical prediction and previously found cointegrating relationships. Moreover, we also find that the change in unemployment rate is positively correlated with the changes in overall crime, burglary, and theft and handling. This result is consistent with its long-run correlation with crimes as well.

As pointed in Chapter Three, one of the weaknesses of this work is the econometric technique we adopted for cointegration test. Our approach, Engle-Granger two step procedure, has the advantage of being straightforward to apply and understand. Its biggest limitation, however, is ignoring the possibility of multiple cointegrating relations when there are more than two variables in the equation. Such limitation could be overcome by Johansen technique. However, in the case of multiple cointegrating relations, such technique will induce the problem of identifying the "true" cointegation between the variables which is rather difficult, if not impossible, to do. To cope with the possibility of multiple cointegrating vectors among our concerned variables, we have carried out both eigenvalue and trace tests to detect the number of cointegration vectors. Our results suggest only one cointegrating relation for each crime rate under analysed. Therefore, as a starting point of our analyses, our adoption of the Engle-Granger two step procedure is appropriate.

The second weakness of this chapter is our rather limited data set. On the one hand, the relatively small sample size is a common issue for annual time series data. However, this fact restricts us from including a relatively complete set of explanatory variables into our analyses and thus expose us under the risk of omitted variable bias. Moreover, time series estimations are only based on temporal variations in the variables. However, the variables employed in our model, such as detection rate and unemployment, are likely to vary across areas. Neglecting area variations could increase the risk of generating biased and misleading results. In order to tackle this issue, we apply panel data analysis in our next chapter.

In Chapter Four, we carry on our analyses by employing panel data disaggregated by 43 police force areas in England and Wales over the period 1992-2005. Such data set not only provides much larger sample size and variations across areas in the variables, it also enables us to eliminate area-specific fixed effects that would be otherwise correlated with independent variables and thus bias our results. We further extend our analyses by incorporating violent crimes into the picture. Therefore, we study six types of crime rates in total: burglary, theft and handling, fraud and forgery, robber, sexual offences, and violence against the person. We formulate our empirical model following Becker (1968) and Ehrlich (1973) and relate each crime rate to crime-specific detection rate, crime-specific prison population, Gini coefficient, unemployment rate, real earnings, as well as the proportion of people aged 15-24. In addition, we also allow crime rate to be affected by its one year lagged value to incorporate the persistence in crime. We adopt generalized method of moments (GMM) technique, in addition to OLS and fixed-effect models, to eliminate area-specific fixed effects. Furthermore, this technique also allows us to apply instrumental variables for the three endogenous variables including detection rate, prison population, as well as the once-lagged crime rate. We choose to use all the available lags starting from the second lags as the instruments of the endogenous variables. As a unique feature, we check the robustness of coefficients not only by applying different regression techniques, such as OLS, fixed-effect and GMM, but also by applying a larger data set (though disaggregated at a somewhat coarser level for some variables) from 1987-2005. The differences of the two data sets have been given detailed description in that chapter.

We summarize the main findings as follows. First, property crimes are explained better by the empirical model than violent crimes. Such result is not surprising because we construct our model by assuming that crime-influential factors affect crime through changing its potential cost and monetary gains. As property crimes aim to obtain financial benefit, they should be more responsive to the variables representing social-economic factors. Second, the law enforcement instruments, represented by detection rate and prison population, are the strongest predictors and show negative effects on each type of crime. This finding has confirmed the predictions made by both Becker (1968) and Ehrlich (1973) that either higher probability of apprehension or more severe punishment will reduce crime through deterrence and incapacitation effects. Third, the social economic factors are represented by Gini coefficient, unemployment and real earnings. All three variables have constantly shown opportunity effects on property crimes. Contrary to expectation, Gini coefficient is negatively correlated with crimes, particularly property crimes, in most cases. We try to explain this finding by arguing that an increase in Gini coefficient could reduce crimes by decreasing the number of rich people and thus the available crime opportunities. However, we still should be cautious with this interpretation because the variable of Gini coefficient is aggregated on national level and thus cannot reflect area variations. Such data limitation could affect the estimation results. Both unemployment and real earnings mainly pick up opportunity effects on property crimes. Given the prediction that both variables could have opposing effects on crime, motivation and opportunity, our results suggest that, during the period being examined, the opportunity effects of both variables are stronger than their motivation effects.

The empirical model and estimation techniques adopted in our panel data analyses have tackled a number of potential issues such as limited sample size, area-specific fixed effects, endogenous law enforcement variables and so on. However, this work can be improved in at least two aspects. On the one hand, the variables of prison population and Gini coefficient are aggregated on national level due to unavailability and show no variation across areas. This data limitation could obstruct the analyses to reveal their true relationships with crime rates. Furthermore, their aggregation level has also made it impossible to eliminate year-specific fixed effects by applying dummy variables. Instead, we have to incorporate a linear time trend in the empirical model to control for area-invariant unobserved factors. On the other hand, this study, like most of other papers in this field, assumes that the crime rate in one area is only affected by its local factors. However, such assumption can be challenged by arguing that offenders are mobile and thus are able to choose where to commit crimes by comparing the relative opportunities between different locations. Neglecting such effect from an empirical model is equivalent to omitting an important explanatory variable from the equation of crime. We address these mentioned problems in the next chapter.

In Chapter Five, we aim to test whether there is spatial spillover effect in crime. We are motivated by the fact that, if offenders are mobile and economically rational, they are able to choose where to commit crimes by comparing the relative crime opportunities and costs of different locations. Tougher law enforcement in one area implies higher cost of committing crimes. Other things being equal, tightened crime control in one area could drive its local offenders spillover into neighbouring areas for their relatively lower cost of committing crimes. Similarly, economically affluent area could attract potential offenders to spillover from neighbouring areas due to its relatively target-rich environment. Therefore, in this chapter, we test the hypothesis that the crime rate in one area depends not only on its local crime-related factors, but also on those factors in neighbouring areas.

We firstly construct a simple theoretical model containing two regions. By assuming individuals are utility maximizing, we show that an offender will spillover into neighbouring area to commit crime if his expected return (measured by potential gain minus expected punishment) from committing crime in neighbouring area netting his travelling cost exceeds his expected return from committing crime in his home area. Accordingly, we predict that the number of crimes due to criminal spillovers from region 1 into region 2 is affected by the relative crime opportunities and expected punishment between the two regions.

In order to test our prediction, we formulate two empirical models. In model one, we allow the crime rate in one area is not only depending on its local explanatory variables including detection rate, unemployment rate, real earnings, and the proportion of young people, but also depending on the crime rates of neighbouring areas. This is because crime rate itself could reflect the availability of crime opportunities. Alternately, when the number of suitable crime targets is exogenously given, higher crime rate could imply more fierce competition between offenders and thus fewer potential opportunities. In model two, we assume that the crime rate of one area is affected by its local explanatory variables as well as those factors from its neighbouring areas.

We apply panel data disaggregated by 43 police force areas in England and Wales covering the period 1998-2001 to estimate our empirical models. The application of panel data in spatial analysis is a significant advance over most other spatial analysis papers. Such data structure enables us to explicitly incorporate both area-specific and year-specific unobserved factors that could be otherwise correlated with the explanatory variables. Our results show that model one performs well to reflect the spatial dependence between crime rates. We have found positive and significant spatial dependence in the crime rates of neighbouring areas for burglary, theft and handling, fraud and forgery and sexual offences. This result supports our expectation that areas with high crime rates have spillover effects on neighbouring areas due to the migration of potential offenders. Among the explanatory variables, the strongest predictors for different crimes are once-lagged detection rate (using lagged detection rate to avoid its endogeneity with crime rates) and real earnings: while the lagged detection rate constantly has negative effect on crimes, real earnings always shows positive impact on property crimes. These results are consistent with predictions as well as our findings from panel data analyses. On the other hand, model 2 performs less well to reflect the spillover effects of explanatory variables. The crime-related variables from neighbouring areas broadly have insignificant coefficients. However, their signs are consistent with expectation in most cases.

Our results from detecting spatial spillovers of crimes imply that it is meaningful to improve the coordination among police force areas. Simply tighten the law enforcement in certain areas will drive the potential offenders spilling over into neighbourhoods. Therefore, increasing the crime control all over the country and equalling the wealth distribution among areas can be effective policies to combat crimes.

6.2 PROSPECTS FOR FUTURE RESEARCHES

Based on the results of model 2 in spatial analyses, we cannot draw the conclusion that crime rate in one area is not affected by the crime-influencing factors in neighbouring areas. Our broadly insignificant results in this model could be caused by the disaggregation level we adopted. Police force areas could still be too big to pick up the spillover effects of explanatory variables, in future research one might try more disaggregated spatial units such as neighbourhoods within a city.

Another angle for future research is motivated by our finding in panel data analysis: the variable of real earnings constantly has positive correlation with property crimes picking up its opportunity effect based on the data 1992-2005. As we apply a longer data 1987-2005, this variable shows negative impact on burglary and fraud, indicating its motivation effect. We try to explain this change by the fact that 1992-2005 is a period of steady economic growth. Under such circumstances, people are more likely to spend and provide more opportunities for property crimes, while potential offenders are less likely to face severe economic difficulties. Therefore, property crimes committed may largely be due to increased opportunities and more attractive targets, rather than economic desperation. Consequently, during economic growth, the opportunity effect of real earnings could be stronger than its motivation effects. On the other hand, the data 1987-2005 includes the economic recession in the early 1990s. During economic downturn, potential offenders are more likely to be facing financial problems and take committing crimes as a solution. A decrease in the earnings would not only reduce the opportunities for property crimes as people spend less, more importantly, it could turn more people into potential offenders due to their economic desperation. Therefore, in economic recession, the motivation effect of real earnings could be stronger than its opportunity effect. In the future, one could try to look at whether economic growth and recession change the correlations between property crimes and social-economic factors.

These outstanding issues and shortcomings we hope will inspire further research in this area.

APPENDICES

Appendix I: Police force areas and local authorities

Police Force Areas	Areas Responsible for by County, District; and Unitary Authority
Avon and Somerset Constabulary	Avon and Somerset; Bath and Northeast Somerset, Bristol, North Somerset, South Gloucestershire
Bedfordshire Police	Bedfordshire; Luton
Cambridgeshire Constabulary	Cambridgeshire; Peterborough
Cheshire Constabulary	Cheshire; Halton, Warrington
City of London Police	; City of London
Cleveland Police	Cleveland; Hartlepool, Middlesbrough, Redcar and Cleveland, Stockton-on-Tees
Cumbria Constabulary	Cumbria
Derbyshire Constabulary	Derbyshire; Derby
Devon and Cornwall Constabulary	Cornwall and Isles of Scilly, Devon; Plymouth, Torbay
Dorset Police	Dorset; Bournemouth, Poole
Durham Constabulary	Durham; Darlington
Dyfed-Powys Police	Carmarthenshire, Ceredigion, Pembrokeshire, Powys,
Essex Police	Essex; Southend-on-Sea, Thurrock
Gloucestershire Constabulary	Gloucestershire;
Greater Manchester Police	Greater Manchester
Gwent Police	Blaenau Gwent, Caerphilly, Monmouthshire, Newport, Torfaen
Hampshire Constabulary	Hampshire and Isle of Wight; Portsmouth, Southampton
Hertfordshire Constabulary	Hertfordshire
Humberside Police	; East Riding of Yorkshire, Kingston upon Hull, Northeast Lincolnshire, North Lincolnshire
Kent Police	Kent; Medway
Lancashire Constabulary	Lancashire; Blackburn, Blackpool
Leicestershire Constabulary	Leicestershire; Leicester, Rutland
Lincolnshire Police	Lincolnshire
Merseyside Police	Merseyside
Metropolitan Police Service	Greater London (excluding City of London)
Norfolk Constabulary	Norfolk
North Wales Police	Conway, Denbighshire, Flintshire, Gwynedd, Isle of Anglesey, Wrexham
North Yorkshire Police	North Yorkshire; York
Northamptonshire Police	Northamptonshire
Northumbria Police	Northumberland, Tyne and Wear
Nottinghamshire Police	Nottinghamshire; Nottingham
South Wales Police	Bridgend, Cardiff, Merthyr Tydfil, Neath Port Talbot, Rhondda Cynon Taff, Swansea, The Vale of Glamorgan
South Yorkshire Police	South Yorkshire
Staffordshire Police	Staffordshire; Stoke-on-Trent
Suffolk Constabulary	Suffolk
Surrey Police	Surrey
Sussex Police	East Sussex, West Sussex; Brighton and Hove
Thames Valley Police	Buckinghamshire, Oxfordshire; Bracknell Forest, Milton Keynes, Reading, Slough, West Berkshire, Windsor and Maidenhead, Wokingham
Warwickshire Police	Warwickshire
West Mercia Police	Worcestershire; Herefordshire, Shropshire, Telford and Wrekin
West Midlands Police	West Midlands
West Yorkshire Police	West Yorkshire
Wiltshire Police	Wiltshire; Swindon

Appendix II: Police force areas and regions in England and Wales

Regions	**Polices Force Areas**
North East, ENGLAND	Cleveland Police, Durham Constabulary, Northumbria Police
North West, ENGLAND	Cheshire Constabulary, Cumbria Constabulary, Greater Manchester Police, Lancashire Constabulary, Merseyside Police
Yorkshire and the Humber, ENGLAND	Humberside Police, North Yorkshire Police, South Yorkshire Police, West Yorkshire Police
East Midlands, ENGLAND	Derbyshire Constabulary, Leicestershire Constabulary, Lincolnshire Police, Northamptonshire Police, Nottinghamshire Police
West Midlands, ENGLAND	Staffordshire Police, Warwickshire Police, West Mercia Police, West Midlands Police
East, ENGLAND	Bedfordshire Police, Cambridgeshire Constabulary, Essex Police, Hertfordshire Constabulary, Norfolk Constabulary, Suffolk Constabulary
London, ENGLAND	City of London Police, Metropolitan Police Service
South East, ENGLAND	Hampshire Constabulary, Kent Police, Surrey Police, Sussex Police, Thames Valley Police
South West, ENGLAND	Avon and Somerset Constabulary, Devon and Cornwall Constabulary, Dorset Police, Gloucestershire Constabulary, Wiltshire Police
WALES	Pyfed-Powys Police, Gwent Police, North Wales Police, South Wales Police

Appendix III: Tables of statistical summaries for independent variables 1987-2005

Table Ap-1 - Statistics for detection rates 1987-2005

		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	31	75	67	33
Median	29	78	73	31
Maximum	69	106	124	96
Minimum	14	24	21	10
Std. Dev.	9	14	21	13
Observations	817	817	688	688
		Property crimes		
		Burglary	Theft and handling	Fraud and forgery
Mean		21	26	52
Median		19	24	50
Maximum		56	54	99
Minimum		7	8	9
Std. Dev.		9	9	18
Observations		817	817	817

Table Ap-2 - Statistics for prison population 1987-2005

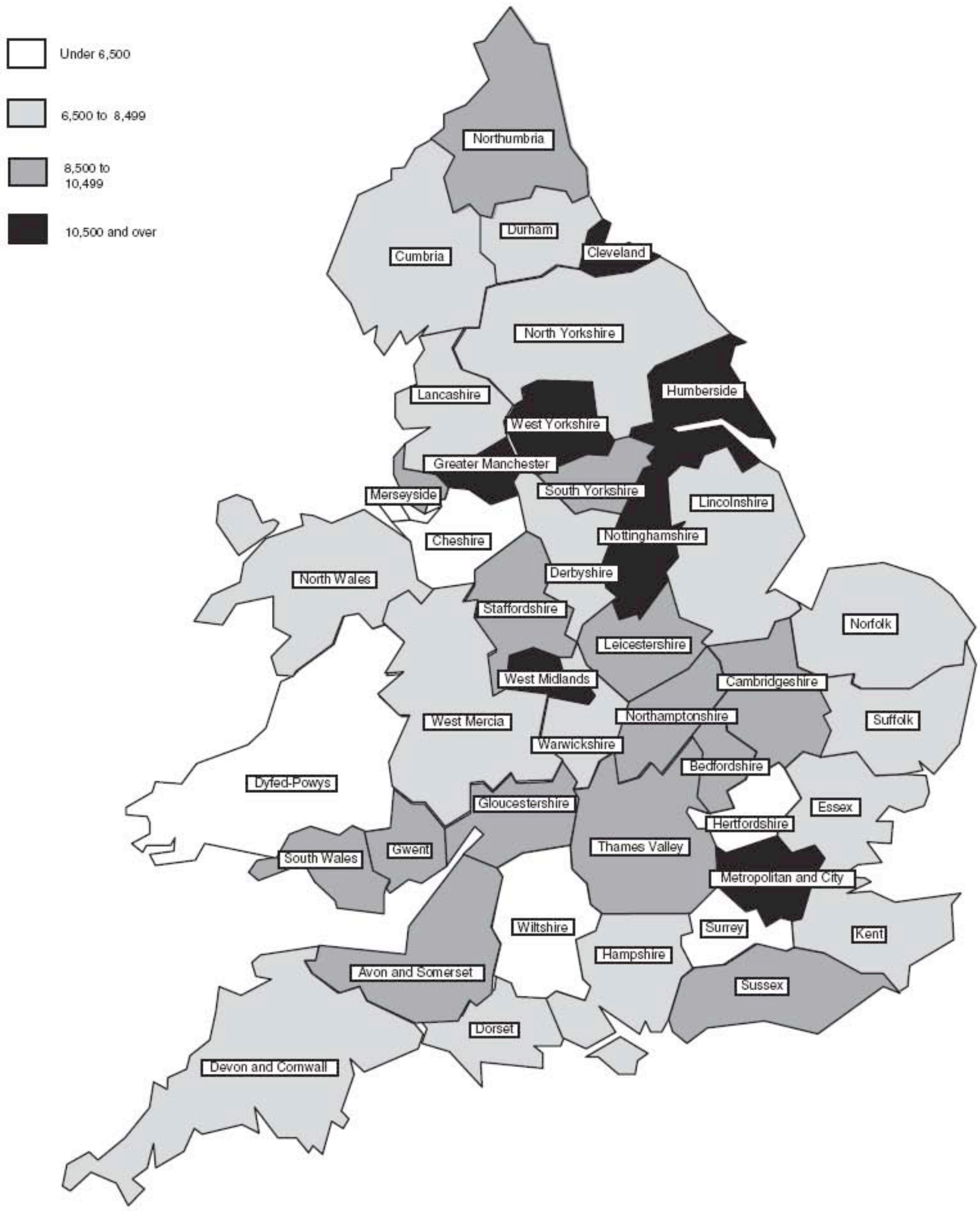
		Violent crimes		
	Overall crime	Violence against person	Sexual offences	Robbery
Mean	88	19	8	11
Median	83	19	8	11
Maximum	116	28	12	16
Minimum	64	14	5	8
Std. Dev.	18	4	2	3
Observations	817	817	817	817
		Property crimes		
		Burglary	Theft and handling	Fraud and forgery
Mean		14	10	10
Median		16	11	11
Maximum		18	14	14
Minimum		9	7	7
Std. Dev.		3	2	2
Observations		817	817	817

Table Ap-3 - Statistics for the other independent variables 1987-2005

	Gini coefficient	Unemployment	Real average weekly earnings	Young people
Mean	38.13	5.82	141.58	12.90
Median	38.00	5.25	137.56	12.81
Maximum	40.40	15.01	452.06	17.41
Minimum	36.10	1.71	27.22	8.11
Std. Dev.	1.20	2.77	82.07	1.55
Observations	817	817	817	817

Appendix IV: Distribution of overall crime rate across police force areas in England and Wales (2000/2001)

Figure 2.5 Recorded crime per 100,000 population by police force area 2000/01



BIBLIOGRAPHY

Anderson, T.G. and Sorenson, B.E. (1996) “GMM estimation of a stochastic volatility model: A Monte Carlo study”, *Journal of Business and Economic Statistics*: 328 – 352

Andresen, Martin A. (2006) ‘A spatial analysis of crime in Vancouver, British Columbia: a synthesis of social disorganization and routine activity theory’, *The Canadian Geographer*, Vol. 50, No. 4: 487-502

Anselin, Luc (1988) *Spatial Econometrics: Methods and Models*, Boston: Kluwer Academic Publishers

____ (1999) ‘Spatial econometrics’ in B. Baltagi ed. *Companion in Theoretical Econometrics*, Oxford: Basil Blackwell

Anselin, L. and Bera, A. (1998) ‘Spatial dependence in linear regression models with an introduction to spatial econometrics’ in A. Ullah and D.E.A. Giles eds., *Handbook of Applied Economic Statistics*, New York: Marcel Dekker

Anselin, Luc; Cohen, Jacqueline; Cook, David; Gorr, Wilpen; and Tita, George (2000) ‘Spatial analyses of crime’, *Criminal Justice*, Vol. 4: 213-262

Becker, Gary S. (1968) “Crime and Punishment: An Economic Approach”, *The Journal of Political Economy*, 76 (2): 169-217

Brantingham, P. L. and Brantingham, P. J. (1982) “Mobility, Notoriety, and Crime: A Study of Crime Patterns in Urban Nodal Points”, *Journal of Environmental Systems*, 11: 89-99

______ (1993) ‘Environment, routine, and situation: toward a pattern theory of crime’ in R.V. Clarke and M. Felson eds, *Routine Activity and Rational Choice: Advances in Criminological Theory*, Vol. 5, New Brunswick, New Jersey: Transaction Publishers

Britt, Chester L. (1994) “Crime and Unemployment among Youths in United States, 1958-1990: A Time Series Analysis”, *American Journal of Economics and Sociology*, 53 (1): 99-109

Cahill, Meagan and Mulligan, Gordon (2003) ‘The determinants of crime in Tucson, Arizona’, *Urban Geography*, No. 24: 582-610

____ (2007) ‘Using geographically weighted regression to explore local crime patterns’, *Social Science Computer Review*, Vol. 25, No. 2, (summer): 174-193

Cantor, David and Land, Kenneth C. (1985) “Unemployment and Crime Rates in the Post-World War II United States: A Theoretical and Empirical Analysis”, *American Sociological Review*, 50 (3): 317-332

Carmichael, Fiona and Ward, Robert (2000) “Youth Unemployment and Crime in the English Regions and Wales”, *Applied Economics*, (32): 559-571

______ (2001) “Male Unemployment and Crime in England and Wales”, *Economic Letters*, (73): 111-115

Choe (2008) “Income inequality and crime in the United States”, *Economics Letters*, (101): 31-33

Cohen, J. and Tita, G. (1999) “Spatial Diffusion in Homicide: An Exploratory Analysis”, *Journal of Quantitative Criminology*, 15 (4): 451-493

Cohen, Lawrence E. and Felson, Marcus (1979) ‘Social change and crime rate trends: a routine activity approach’, *American Sociological Review*, No.44: 588-605

Cohen, Lawrence E., Felson, Marcus, and Land, Kenneth C. (1980) “Property Crime Rates in the United States: A Macrodynamic Analysis, 1947-1977: with Ex Ante Forecasts for the Mid-1980s”, *The American Journal of Sociology,* 86 (1): 90-118

Cornish, L. E. and Clarke, R. V. (1986) *The Seasoning Criminal: Rational Choice Perspectives on Offending,* New York: Springer-Verlag

Curry, David and Spergel, Irving (1988) “Gang Homicide, Delinquency, and Community”, *Criminology*, 26: 381-405

Dickey, D. A. and Fuller, W. A. (1979) “Distributions of the Estimators for Autoregressive Time Series with A Unit Root”, *Journal of the American Statistical Association*, 74: 427-431

Doyle, Joanne M., Ahmed, Ehsan and Horn, Robert N. (1999) "The Effects of Labor Markets and Income Inequality on Crime: Evidence from Panel Data", *Southern Economic Journal*, 65 (4): 717-738

Edmark, Karin (2003) "The Effects of Unemployment on Property Crime: Evidence from A Period of Unusually Large Swings in the Business Cycle", *Uppsala University, Department of Economics,* Working paper series 2003: 14

Ehrlich, Isaac (1972) "The Deterrent Effect of Criminal Law Enforcement", *The Journal of Legal Studies,* 1 (2): 259-276

______ (1973) "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation", *The Journal of Political Economy,* 81 (3): 521-565

______ (1996) "Crime Punishment, and the Market of Offenses", *The Journal of Economic Perspectives*, 10 (1): 43-67

Engle, Robert F. and Granger, C. W. J. (1987) "Co-Integration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55 (2): 251-276

Engstad, P. A. (1975) "Environmental Opportunities and the Ecology of Crime", in Silverman, R. A. and Teevan, J. J. (ed.) Toronto: Butterworths

Fabrikant, Richard (1979) 'The distribution of criminal offenses in an urban environment: a spatial analysis of criminal spillovers and of juvenile offenders', *American Journal of Economics and Sociology*, Vol. 38, No. 1 (January): 31-47

Felson, Marcus (1986) 'Linking criminal choices, routine activities, informal crime control, and criminal outcomes', in D. Cornish and R.V. Clarke eds, *The Seasoning Criminal: Rational Choice Perspectives on Offending*, New York: Springer-Verlag

____ (1994) *Crime and Everyday Life: Impact and Implications for Society,* Thousand Oaks, California: Pine Forge Press

Fleisher, Belton M. (1963) "The Effect of Unemployment on Juvenile Delinquency", *The Journal of Political Economy*, 71 (6): 543-555

Fougère, Denis; Kramarz, Francis and Pouget, Julien (2003) "Crime and Unemployment in France", *Bonn: Institute for the Study of Labor*, IZA Discussion papers 2009

Gould, Eric D., Weinberg, Bruce A. and Mustard, David B. (2002) "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997", *The Review of Economics and Statistics*, 84 (1): 45-61

Greenberg, David F. (2001) "Time Series Analysis of Crime Rates", *Journal of Quantitative Criminology*, 17 (4): 291-327

Gylys, Julius A. (1974) "The Interdependence of Municipal and County Police Forces", *American Journal of Economics and Sociology*, 33: 75-88

Hale, Chris (1998) "Crime and the Business Cycle in Post-War Britain Revisited", *British Journal of Criminology,* 38 (4): 681-698

Hale, Chris and Sabbagh, Dima (1991) "Testing the Relationship between Unemployment and Crime: A Methodological Comment and Empirical Analysis Using Time Series Data from England and Wales", *Journal of Research in Crime and Delinquency,* 28: 400-417

Harries, Richard (2003) "Modelling and Predicting Recorded Property Crime Trends in England and Wales—A Retrospective", *International Journal of Forecasting,* 19: 557-566

Harris, K. D. (1995) 'The ecology of homicide and assault: Baltimore city and country, 1989-1991', *Studies in Crime and Crime Prevention*, No.4: 44-60

Katz, Lawrence, Levitt, Steven D. and Shustorovich, Ellen (2003) "Prison Conditions, Capital Punishment, and Deterrence", *American Law and Economics Review,* 5 (2): 318-343

Kelejian, H. H. and Prucha, I. (1998) 'A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances', *Journal of Real Estate Finance and Economics*, No. 17: 99-121

Kelejian, H. H. and Robinson, D. P. (1993) 'A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model', *Papers in Regional Science*, No. 72: 297-312

Kelly (2000) “Inequality and Crime”, *The Review of Economics and Statistics,* 82 (4): 530 – 539

Land, K. and Deane, G. (1992) ‘On the large-sample estimation of regression models with spatial- or network-effects terms: a two stage least squares approach’ in P. Marsden ed. *Sociological Methodology,* San Francisco: Jossey-Bass

Levitt, Steven D. (1996) “The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation”, *The Quarterly Journal of Economics,* 111 (2): 319-351

______ (1998a) “Juvenile Crime and Punishment”, *Journal of Political Economy,* 106 (6): 1156-1183

______ (1998b) “Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?”, *Economic Enquiry,* 36 (3): 353-372

______ (1999) “The Limited Role of Changing Age Structure in Explaining Aggregate Crime Rates”, *Criminology,* 37 (3): 581-597

______ (2001) “Alternative Strategies for Identifying the Link between Unemployment and Crime”, *Journal of Quantitative Criminology,* 17 (4): 377-390

Machin, Stephen and Meghir, Coastas (2000) “Crime and Economic Incentives”, *London: The Institute for Fiscal Studies,* Working papers W00/17

MacKinnon, J. G. (1996) "Numerical Distribution Functions for Unit Root and Cointegration Tests", *Journal of Applied Econometrics,* 11 (6): 601-618

Malczewski, J. and Poetz, A. (2005) 'Residential burglaries and neighbourhood socioeconomic context in London, Ontario: global and local regression analysis', *The Professional Geographer,* Vol. 57, No. 4: 516-529

Martin, David (2002) 'Spatial patterns in residential burglary: assessing the effects of neighbourhood social capital', *Journal of Contemporary Criminal Justice*, Vol. 18, No. 2: 132-146

Mehay, Stephen L. (1977) "Interjurisdictional Spillovers of Urban Police Services", *Southern Economic Journal,* 43: 1352-1359

Newman, Oscar (1972) *Defensible Space*, New York: Macmillan

Phillips, P. C. B. and Perron, P. (1988) "Testing for a Unit Root in Time Series Regression", *Biometrika,* 75: 335-346

Press, James S. (1971) *Some Effects of An Increase in Police Manpower in the 20th Precinct of New York City,* New York: The New York City Rand Institute

Pyle, D. J. and Deadman, D. F. (1994a) "Crime and the Business Cycle in Post-War Britain", *British Journal of Criminology,* 34 (3): 339-357

______ (1994b) “Crime and Unemployment in Scotland: Some Results”, *Scottish Journal of Political Economy,* 41 (3): 314-324

_______ (1997) “Forecasting Recorded Property Crime Using A Time-Series Econometric Model”, *British Journal of Criminology,* 37 (3): 437-445

Reilly, Barry and Witt, Robert (1992) “Crime and Unemployment in Scotland: An Econometric Analysis Using Regional Data”, *Scottish Journal of Political Economy,* 39 (2): 213-228

______ (1996) “Crime, Deterrence and Unemployment in England and Wales: An Empirical Analysis”, *Bulletin of Economic Research,* 48 (2): 137-159

Raphael, Steven and Winter-Ebmer, Rudolf (2001) “Identifying the Effect of Unemployment on Crime”, *Journal of Law and Economics*, 44 (1): 259-283

Roodman (2009) “How to use xtabound2: An Introduction to difference and system GMM in Stata”, *The Stata Journal,* 9 (1): 86 – 136

Roncek, D.W. and Bell, R. (1981) ‘Bars, blokes and crime’, *Journal of Environmental Systems,* No.11: 35-47

Saridakis, George (2008) “Violent Crime and Incentives in the Long-Run: Evidence from England and Wales”, *Loughborough University Business School Research Series Papers,* 2008 (2)

Saridakis, George and Spengler, Hannes (2009) “Crime, Deterrence and Unemployment in Greece: A Panel Data Approach”, *Berlin: German Institute for Economic Research,* Discussion papers 853

Scorzafave and Soares (2009) “Income inequality and pecuniary crimes”, *Economics Letters,* (104): 40-42

Shaw, Clifford R. and McKay, Henry D. (1942) *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquents in Relation to Differential Characteristics of Local Communities in American Cities*, Chicago: University of Chicago Press

Tabarrok, Alexander and Helland, Eric (2009) “Measuring Criminal Spillovers: Evidence from Three Strikes”, *Review of Law and Economics*, 5 (1)

Witt, Robert; Clarke, Alan and Fielding, Nigel (1998) “Crime, Earnings Inequality and Unemployment in England and Wales”, *Applied Economics Letters*, (5): 265-267

______ (1999) “Crime and Economic Activity: A Panel Data Approach”, *The British Journal of Criminology,* 39 (3): 391-400

Wolpin, Kenneth I. (1978) “An Economic Analysis of Crime and Punishment in England and Wales, 1894-1967”, *The Journal of Political Economy*, 86 (5): 815-840