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.

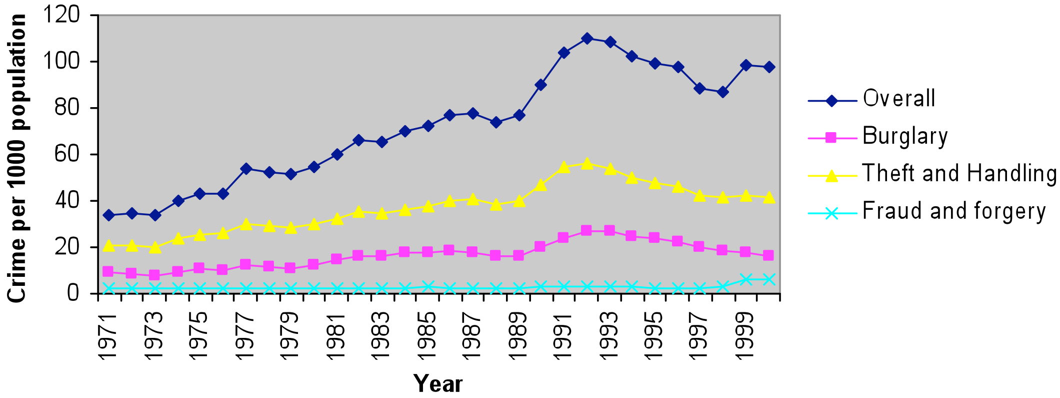
Given the potential shortcomings of time series data as mentioned above, we argue that applying time series data and techniques is still meaningful for investigating the correlation between crime and its relevant factors. In this chapter, we are particularly interested in using national-level time series 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.

Crime rates in England and Wales 1971-2000



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:

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

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

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

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

	ADF		PP	
Variable	Level	First Difference	Level	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

	ADF		PP	
Variable	Level	First Difference	Level	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)

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

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

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.