Crime and Economic Incentives

December 1998 – First Draft
September 2000 – This Draft

Stephen Machin* **and Costas Meghir****

* *Department of Economics, University College London, Institute for Fiscal Studies and Centre for Economic Performance, London School of Economics.*

** *Department of Economics, University College London, Institute for Fiscal Studies and CEPR*

Abstract

We explore the role that economic incentives, particularly changes in wages at the bottom end of the wage distribution, play in determining crime rates. We use data on the police force areas of England and Wales between 1975 and 1996. We find that falls in the wages of unskilled workers leads to increases in crime. We carry out a number of experiments with different wage measures, including a wage measure that accounts for the effects of changes in the composition of employment. These reinforce the picture of a strong impact of wages on crime. The result that incentives play a central role is reinforced further by the strong impact on crime of deterrence measures and of a measure of the returns to crime.

Acknowledgements

We thank Gary Becker, Jeff Grogger, Richard Harries, Steve Levitt, Casey Mulligan, Steve Raphael, Rudolf Winter-Ebmer and participants in the 1999 American Economic Association meeting in New York, the 1999 American Society of Criminologists meeting in Toronto and the 2000 British Society of Criminology conference in Leicester for useful and constructive comments. We would also like to thank Steve Gibbons, Lupin Rahman and Ben Klemens for excellent research assistance.

I. Introduction

In the United Kingdom crime increased rapidly through the 1970s and 1980s, reaching record levels by the start of the 1990s, becoming an important public policy issue. The increase was far greater than that experienced in the rest of Europe. Moreover, property crime rose very rapidly in the UK through the 1970s, 80s and early 90s. Indeed, a recent US Bureau of Justice Statistics study (Langan and Farrington, 1998) states the level of property crime to be higher in 1996 in England and Wales than in the US.[1] At the same time, and in contrast to the rest of Europe, wage inequality, reached record highs this century at the start of the 1990s (Machin, 1996, 1998). Central to this increase has been the rapidly deteriorating labour market position of less skilled workers in the UK at the bottom end of the wage distribution.[2]

Simple economic models predict that declining labour market opportunities are likely to alter the incentives for individuals to participate in legitimate (labour market) or illegitimate (criminal) activities. In this paper we investigate this in some detail, by examining the extent to which rising crime is related to the worsening labour market position of less skilled workers using area-level data from England and Wales from the mid-1970s to the mid-1990s.

Because the social and economic costs associated with high crime rates are potentially enormous, this issue has probably not been examined in enough detail, particularly in the UK. There are existing findings that sometimes point to significant correlations between crime and unemployment (e.g. Reilly and Witt, 1996), but the link with changes in actual labour market

[1] For example, Langan and Farrington report that in 1995 (the most recent year for which comparable data is available) there were 7.6 robberies and 82.9 burglaries per 1000 households in England and Wales as compared to 5.5 and 47.5 per 1000 households in the US. Of course, comparisons of violent non-property crimes show the US to have much higher rates (e.g. the murder rate is around 6 times higher in the US). See Freeman (1996) for the public policy issues relating to crime in the US.

[2] A large literature now exists on what may have been behind the adverse demand shifts against the less skilled. This work has focussed on the role of skill-biased technological changes (e.g. the rapid introduction of computers to the workplace) and rising international competition (mainly in the form of increased wage competition with low-wage

opportunities (e.g. measured in terms of wages) has been much less explored.[3] This is true despite the fact that increases in crime have been given a lot of attention by several different social science disciplines over the years.

The structure of the paper is as follows. In Section 2 we present a simple structural model of links between crime and the labour market, from which we derive our estimating equations. In Section 3 we describe the data, which covers the police force areas of England and Wales between 1975 and 1996. Section 4 presents the estimated crime models, while Section 5 concludes.

II. The Economic Model

The data we will be using is aggregated police force area data.[4] In order to help interpret the results we start by setting up a simple choice theoretic model of crime participation and labour supply. The basic set up of the model is not unlike that of Ehrlich (1973) who formulates a model where individuals consider whether to engage in legal or illegal activities by comparing the value of earnings from crime net of the loss of earnings from not working, taking into account the probability that they are apprehended. In the Ehrlich model crime and work are substitutes as each takes time and produces income. We adopt a more general approach for deriving our empirical model where individuals choose between work and crime, but where they may also combine them both. Indeed, this latter point is potentially very important as a number of authors (Grogger, 1998,

---

developing countries). See the survey pieces by Katz and Autor (1999) and Johnson and Stafford (1999) for more details on this work.

[3] But see Grogger (1998), Levitt (1998), Gould, Weinberg and Mustard (1998) and May (1999) for US evidence that suggests wages (especially for those people whose wages are low enough to put them on the margins of committing crimes) are an important determinant of crime.

[4] We are basically forced to use area data as individual data on crime is not available over time for representative samples in the UK. There are occasional surveys of individuals (like the MORI 'Young people and crime survey' of people aged 14-26 in 1992 used in Hansen, 1999) and there are victimisation surveys over time, but none that can be put together to form a coherent picture of what has been happening over time. This is clearly unlike the US case where some surveys do have individual data on crime over time.

and Fagan and Freeman, 1997, among others) stress that many crimes are committed by the employed.

The basic feature of the model is that individuals choose between engaging or not in criminal activity and working or not. We take these decisions to be discrete. Denote the crime indicator by *b* (for bad) and the work participation indicator by *l*. The value of engaging in criminal activity and being successful and working at the same time is then V^{bl}. Normalise the value of being unsuccessful to zero and assume it is the same whether one is working or not. The value of engaging in successful criminal activity while not working is V^{b0}. Finally, the values of not engaging in criminal activity while working is V^{0l} and while not working is V^{00}.

One should notice that here we ignore any intertemporal features of such decisions. As far as the employment decision is concerned this does not appear unrealistic: the growth rate of wages due to experience seems to be very low, if not zero, for low skill workers once they participate in the adult labour market (see Meghir and Whitehouse, 1996 and Gosling, Machin and Meghir, 1999).[5] We have little knowledge on the nature of intertemporal links for criminal activity. However, it is likely that a lot of the persistence is due to unobserved heterogeneity and neighbourhood effects (see Glaeser, Sacerdote and Scheinkman, 1996). This is incorporated in the modelling strategy we consider. As is the case with "external habits" (Campbell and Cochrane, 1999) the individual ignores such persistence in his/her decision making, since they are external to his/her behaviour.

We assume that individuals choose the outcome with the highest expected value. Uncertainty in the model originates from the probability of being caught. We denote the probability of being arrested for individual i in year t by P_{it} and this will depend on police resources and other

[5] Of course wage growth can sometimes be rapid in the teenage years, even for the unskilled, but it seems that, by the time they are into their twenties, the wage returns to experience hardly rise at all over the life cycle for the least skilled.

area characteristics. A successful criminal act in area a in year t has a net return r_{at}. When working individual i can earn a wage w_{it} and the low-income unemployed receive transfers T_{it}.[6]

The literature has emphasized peer group effects as being a very important determinant of crime. We follow two approaches to allow for this. First, by including area effects we are likely to capture most of the peer effects that do not fluctuate much from year to year.[7] This will also allow for differences in types of area induced by the differences in the opportunities for crime such as differences between urban and rural areas as emphasized by Glaeser and Sacerdote (1999). We also report some results where we include the lagged area crime rate $\overline{b}_{a,t-1}$ as a proxy for fluctuating peer effects: If peer effects are important, a surge in crime will persist for some time and the lag in crime rate will capture the peer effect.

Given this discussion we specify the (expected) values for each combination of activities in the following way :

i) The value of engaging in crime and being employed:

$$EV^{bl} = [d_1^{bl} r_{at} + d_2^{bl} \log(w_{it}) + d_3^{bl} \overline{b}_{a,t-1} + \text{area} + \text{time} + e_{it}^{bl}](1 - P_{it}) = x_{it}^{'} d^{bl} + v_{it}^{bl}$$

ii) The value of engaging in crime and being unemployed:

$$EV^{b0} = [d_1^{b0} r_{at} + d_2^{b0} T_{it} + d_3^{b0} \overline{b}_{a,t-1} + \text{area} + \text{time} + e_{it}^{b0}](1 - P_{it}) = x_{it}^{'} d^{b0} + v_{it}^{b0}$$

iii) The value of not engaging in crime and being employed:

$$EV^{0l} = [d_1^{0l} \log(w_{it}) + \text{area} + \text{time} + e_{it}^{0l}] = x_{it}^{'} d^{0l} + v_{it}^{0l}$$

iv) the value of being unemployed without engaging in criminal activity:

$$EV^{00} = [d_1^{00} T_{it} + \text{area} + \text{time} + e_{it}^{00}] = x_{it}^{'} d^{00} + v_{it}^{00}$$

---

[6] Note that in the UK earnings related transfers (benefits) are universally available, irrespective of age or marital status, with few restrictions only.
[7] Our empirical work also controls for demographics through the population share of young people.

where e and v represent unobserved heterogeneity reflecting individual tastes for each option.

In all of the above, whenever relevant, the optimal level of criminal activity or of labour supply effort has been substituted out and hence the values depend on the prices/returns from these activities rather than effort levels. According to this simple model engaging in criminal activity will depend on the returns to crime. It will also depend on the returns to working, since these two activities are competing for the individual's leisure time and on the benefit system which determines transfers to low income or unemployed individuals. Transfers matter because they act as an income effect and they affect the returns to work relative to being out of work. The direction of the effects, although theoretically ambiguous, can be derived if we assume that leisure is a normal good and that criminal activity is not enjoyable. In terms of the structural model the effect of transfers will depend on the relative size of d_1^{00} and d_2^{b0}. If transfers have a stronger effect on the probability of being unemployed, and if the leisure costs of crime are low, transfers could increase crime. Unfortunately there is practically no regional variation in the level of benefits, and hence we are unable to shed any light on this issue in our empirical work. Moreover, from a policy perspective, an evaluation of the impact of transfers would need to take into account any general equilibrium effect on wages.

On the other hand we would expect that increased wages reduce crime by shifting the relative incentives associated with participating in legal versus illegal activities. Since these values represent indirect utilities the increase in wages can be thought of as inducing an increase in participation and in hours of work, and possibly displacing criminal activity. However, the possibility of engaging in both activities, implies that the effect of wages is in principle ambiguous and depends on the relative size of the wage effects in the values of engaging in work only and in engaging both in work and crime. Nevertheless, since crime involves effort and the possibility of being caught, it is reasonable to expect that an increase in individual earnings capacity will displace criminal activity.

The factors determining the probability of being caught will have an unambiguous negative effect of reducing crime, according to the model. This is because increases in this probability decrease the *ex ante* return to crime. In most of our analysis, we control for such factors using the conviction rate in each areas (as well as by the area and time effects). Hence increases in the conviction rate should decrease the crime rate. We do discuss the issue of possible reverse causality below.

III. The Empirical Model

Derivation of the Empirical Model

To fix ideas on the way we approach the problem empirically, note that in our data we do not know whether an individual is working or not. Moreover, the aggregate area-level data reports information on the regional employment rates but these are obviously not broken down by criminal activity. Our strategy is to model the aggregate probability of engaging in criminal activity. We derive an expression for this assuming that the unobservables in the value functions are distributed as extreme value.[8] This implies a logistic probability of taking up each of the four options. Adding the probabilities involving committing crimes we then obtain the probability of engaging in criminal activity whether working or not. This is given (dropping the 'it' subscripts) by:

$$C^b = \frac{e^{x'd^{b1}} + e^{x'd^{b0}}}{e^{x'd^{b1}} + e^{x'd^{b0}} + e^{x'd^{01}} + e^{x'd^{00}}}$$

Aggregation at the regional level is helped if we take the log odds ratio, which we then approximate by the linear function:

$$\log[C^b_{at}/(1-C^b_{at})] = d_1 r_{at} + d_2 \log(w_{at}) + d_3 T_{at} + d_4 \overline{b}_{a,t-1} + d_5 \log(P_{at}) + \text{area} + \text{time} + u_{at}$$

where the subscript a denotes area so that C^b_{at} is the area crime rate for area a in year t. The variable $\overline{b}_{a,t-1}$ is a lagged dependent variable. As noted earlier, we assume that peer effects on criminal activity are partly captured by the permanent area effects and partly by the criminal activity in the previous period – hence the lagged dependent variable. However, we will be experiment with and without this variable and we show that the overall conclusion of the impact on incentives are not sensitive to this. The stochastic error term u_{at} reflects unobserved area characteristics that vary over time, implicitly included among the regressors up to now.

As we will describe below our data consists of a police force area panel spanning 20 years. This allows us to estimate area fixed effects (within groups) models to control for time invariant area effects. One important issue here is the possible endogeneity of the conviction rate. It is possible that the conviction rate declines because of an increase of criminal activity when, for example, police force resources remain constant. Such an event would lead to a spurious negative relationship between crime rates and conviction rates. We address this in a number of ways. First, we also use lagged conviction rates. Second, we instrument the current conviction rate with the lag and with a police resources measure. Thus the within groups estimation procedure we use is in certain occasions combined with instrumental variables. Third, we have also looked at average sentence lengths as an alternative measure of the strength of deterrence of the criminal justice system.[9]

---

[8] Strictly speaking the errors are heteroskedastic since they are multiplied by the conviction probability. Given the nature of our data we cannot address this subtle point directly, except by checking the robustness of our results to various specifications changes.

[9] Note that average sentence length may also be thought of as endogenous. However, in this case the likely bias is reversed: as crime rates increase, judges may impose stiffer sentences to send a message to criminals. In this case a positive correlation is induced between crime and sentence length, leading to a downward bias of the coefficient.

Before turning to the wage measure we consider one more point about the empirical implications of the model needs to be made. Quite often crime equations in the literature include the unemployment rate. However, in this model the unemployment rate would simply be a reflection of one aspect of labour supply and is controlled for by the wage rate, which reflects the returns from work and controls for the employment decisions made by individuals. In addition, controlling for unemployment instead of wages misses an important point of the model: Crime and work may co-exist; in a low wage regime labour supply (or work effort) may be reduced and more time (and/or effort) put into crime. This may not show up as decreased labour market participation. So, at least in a competitive labour market the wage rate reflects better the incentives than just unemployment. Nevertheless, if institutional factors prevent wage adjustments at the bottom of the wage distribution wage fluctuations will not fully reflect employment opportunities. For instance, suppose minimum wages are present; in this case some of the unemployment is "involuntary" in the sense that some of the unemployed would be willing to work at their marginal product but are prevented. This could be a concern, but it seems unlikely in the UK since, over the period we are considering, there was no national minimum wage in operation in the UK labour market. The main institutional factor preventing wages from falling too much is the income support system. In this case however, the labour market opportunities are reflected in wages, once we take care of the selection issue due to changes in the composition of the population of workers as they move in and out of employment.

The Wage Measures

We use a variety of wage measures to reflect the benefits from working. Our data on wages is the administrative "New Earnings Survey" which allows us to observe a sample of wages in each police force area. We do not know the skills or education level of criminals or potential criminals. However, it is reasonable to believe that they originate from the bottom of the skill distribution. Thus we present results based on measures of the low wage labour market. First, we use the 25th percentile of the wage distribution in each area. Second, we focus on the wage distribution for the retail trade sector – an industry where low skill workers can easily move in and out of work. [10]

One possible flaw in our wage measures is that some of the fluctuation over time and regions may be induced by differences in employment rates, and hence by differences in the composition of the workforce. In periods of low labour demand, wages would tend to fall and the least productive workers would tend to drop out of the labour force (see Heckman and Sedlacek, 1985, and Blundell, Stoker and Reed, 1999, who show the empirical importance of this in the USA and UK). Observed wages would be an overestimate of the actual wages. In these periods we would also expect criminal activity to increase. Since observed changes in criminal activity will be attributed to lower fluctuation of wages, ignoring this counter-cyclical bias of wages could lead us to overestimate the impact of wages on criminal activity.

Quantiles of the distribution however can be bounded[11], and hence we can carry some sensitivity analysis surrounding this issue. In particular, suppose we make the extreme assumption that the wages that non-workers would earn are all below the wages of those currently working. On the other extreme we can assume that workers are randomly selected with respect to their earnings ability and hence the quantiles of the distribution of wages for those observed working are the same as the quantiles of the unconditional wage distribution. Then the qth quantile of the wage distribution for all workers (currently working or not) in region a and period t (w_{at}^{q}) is bounded by

[10] We utilise the 25th percentile wage because, when we get to the smaller samples in the retail trade wage distributions, we are concerned about possible measurement errors in wages at lower centiles of the distribution.

[11] See Manski (1994). For a recent application of some of these ideas for wages see Blundell, Gosling, Ichimura and Meghir (2000).

(C) $$w_{at}^{qo} \geq w_{at}^{q} \geq F_{at}^{-1}\left(\frac{q - pr(I=0 \mid a,t)}{pr(I=1 \mid a,t)}\right)$$

where w_{at}^{qo} is the qth quantile of the observed wage distribution, $F_{at}^{-1}(k)$ is the point in the observed regional wage distribution in period t corresponding to a proportion of workers k earning below that point.[12] Finally, $pr(I=1 \mid a,t)$ is the proportion of workers in region a and period t and $pr(I=0 \mid a,t) = 1 - pr(I=1 \mid a,t)$. For example, if we are looking for the upper bound on the 25th percentile of the unconditional wage distribution this is bounded below by the $(0.25 - \mathrm{pr(I=0)})/\mathrm{pr(I=1)}$ quantile of the distribution of wages for those observed working. This bound will vary over time and areas both as a result of differences in the employment rates and differences in the observed wage distribution. The lower bound will be satisfied with equality if the market wage of those out of work is always below the market wage of those in work. On the other hand, if we assume that non-workers are a random sample of workers then $w_{at}^{qo} = w_{at}^{q}$.[13]

The above provides us with two different ways of computing the wage measure, reflecting different assumptions about the employment process. Finally, it is worth noting that if all employment fluctuations can be expressed as an additive area effect and a time effect (i.e. no differential changes across areas in the employment rates) including time and area effects in the crime equation will mitigate the requirement to carry out such corrections.

[12] Thus F_{at} is the distribution of wages for workers in area *a* in period t.

[13] Such an assumption is not that unreasonable if we think that sufficient numbers of individuals quit to search for better jobs from the top part of the distribution. One could of course widen the bound to allow for the possibility that *all* non-workers are selected from the top of the wage distribution. We did not attempt this. However, this would lead to an increased estimate of the incentive effects of wages on crime.

III. The Data

Crime Data

Area level crime data is available over time for the 43 police force areas of England and Wales. We have constructed a panel of police force area data for 42 areas (dropping the City of London[14]), beginning in 1975. The crime rate upon which we focus comes from the crime statistics recorded by the police, which are compiled by the Home Office and published annually in *Criminal Statistics*. The measures of crime we use are the numbers of recorded offences of crime against the property and vehicle crime.[15] In terms of our model, this is important since the motivation for more violent crime (e.g. those against the person) may be less related to economic incentives.

In the police force area data, crime against the property is broken down into burglary and theft and handling[16] (exact definitions of the crimes included in the notifiable offences in *Criminal Statistics* for the measures we use and other crime categories are given in Appendix Table A1). For the basic models we have simply added these together to eliminate some of the variation found from area to area based on the differences in individual police officer classification of crime. We do, however, also look at them separately. We also look at vehicle crimes separately in some of our analysis.

---

[14] We exclude the City of London because (a) hardly anyone lives there so its crime rate is artificially high; and (b) there was a massive increase in police resources to combat terrorist activity. Since all our econometric estimates are weighted by area population, the exclusion of the City has no impact at all on our results.

[15] The other possibility is to use victimization data from the British Crime Surveys (BCS) of 1984, 1988, 1992, 1994 and 1996. For the purpose of this study the official recorded crime statistics are the most useful, as they are available based on police force area, to which we are able to match labour market data, and they cover a longer time period on an annual basis. It should, however, be noted that some people argue that the BCS yields a better and preferred measure of actual crime. However, as it is a sample survey, it too potentially suffers from measurement error.

[16] Robbery might also be included. We do not include it (although we did in an earlier version of the paper with very similar results) because the numbers recorded under robbery are small and because some robberies are much more likely to involve violence.

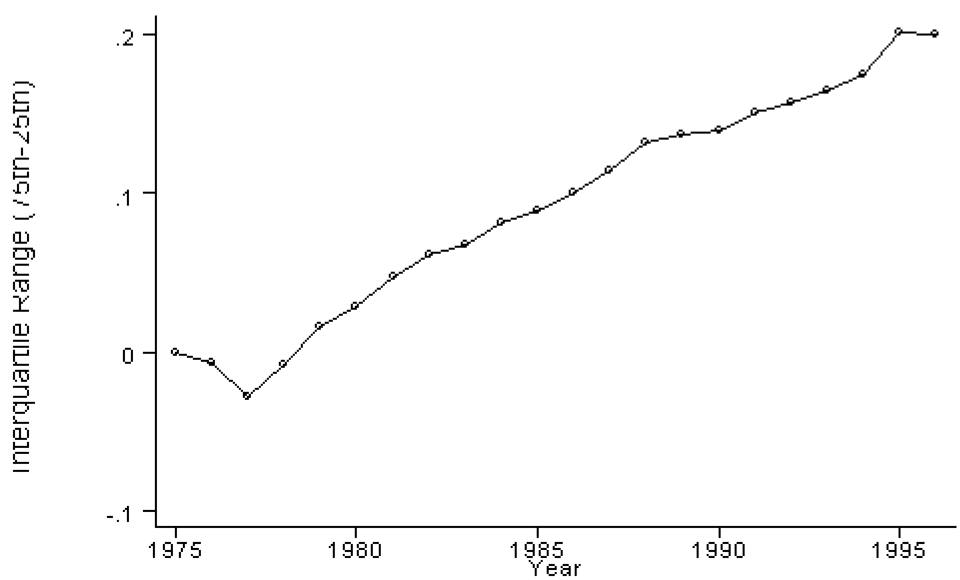
Figure 1 reports what has happened to the property crime rate, to its components and to the vehicle crime rate between 1975 and 1996. There is a clear rise in the property crime rate, from around 38 crimes per 1000 people (.038) in 1975 up to around 68 crimes per 1000 (.068) population by 1996. Looking at the separate components reveals rather similar scale increases for theft and burglary. The same pattern, with a larger percentage rise compared to the initial level, occurs for vehicle crime, which rises from 10 crimes per 1000 to 25 crimes per 1000 between 1975 and 1996. Notice also the dip down in crime from 1992 onwards, revealing a peak in the crime rate in that year, at 80 property crimes and 30 vehicle crimes per 1000 people in England and Wales.

Figure 1: Property Crime Rates per person in England and Wales, 1974-96

Labour Market Data

Area level data on wages is sparse in the UK, but the New Earnings Survey (NES) does contain area-level codes, mainly at county-level, that can be matched to police force areas. From the NES microdata (which is available from 1975 onwards) we have matched data on wages at different points in the police-force area-specific wage distribution to our panel of crime rates. It is well known that, at the same time as crime has risen rapidly, the wage distribution has widened out considerably. Figure 2 shows the interquartile range of the log(wage) distribution by year for England and Wales between 1975 and 1996.[17] As is well known in the wider UK context (Gosling, Machin and Meghir, 2000, Machin, 1996, 1998) the wage distribution became much more unequal from the late 1970s onwards and this pattern is reconfirmed in the Figure.

Figure 2: Changes in the Interquartile Range of the England and Wales Earnings Distribution, 1975-96



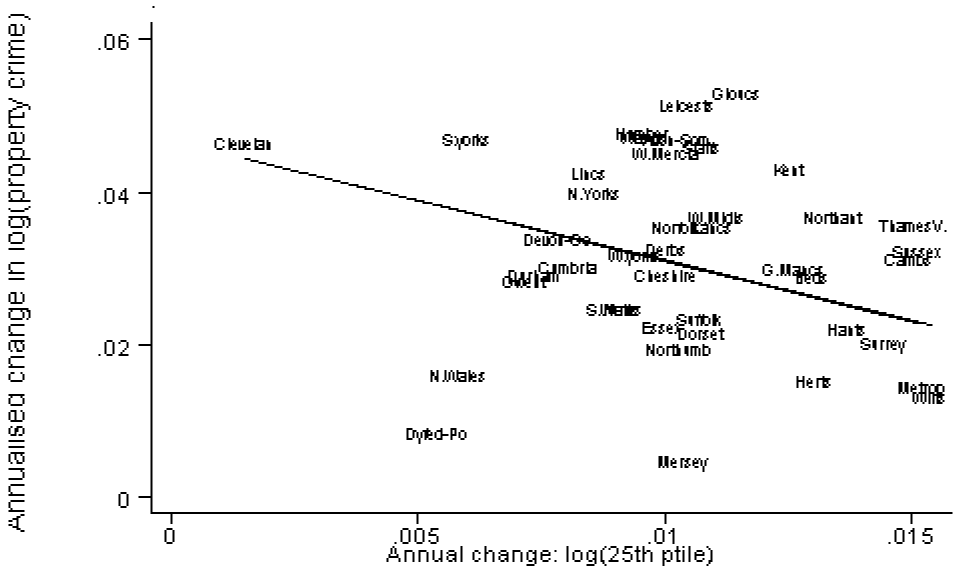
Notes:
1. Source: New Earnings Survey micro-data for all workers in England and Wales.

This rapid rise in wage inequality has led some authors (e.g. Witt, Clarke and Fielding, 1998) to postulate that increasing crime may be linked to rising wage inequality. In our model the prediction is clearer: it is the alteration of labour market opportunities via wages that matters. Increased wage inequality may matter, in that it may be related to the returns from crime. The practical issue that emerges then concerns the appropriate wage measure. As already noted we use measures that we think are appropriate to the crime-work choice (the 25th percentile of the overall wage distribution and the 25th percentile of the retail trade wage distribution). The main idea here is that we are attempting to capture the employment/wage opportunities for the low skill labour market, which is likely to be the relevant one for the bulk of those committing property crimes (e.g. because potential criminals are likely to have low education and labour market attachment).

Figure 3 illustrates the potentially important impact of wages on property crime rates. The Figure plots the 1975-96 change in log(property crime) against the 1975-96 change in the log(25th percentile wage) across police force areas (the numbers are annualised averages). The fitted line is a weighted least squares regression (weights being the area population). There is a very clear, statistically significant, negative association. Areas with lower wage growth at the bottom end of the wage distribution experienced faster rising property crime rates between the mid-1970s and mid-1990s. The econometric results that follow in Section 4 subject this basic finding to a more rigorous examination in more detailed econometric models.

[17] These numbers are taken from the NES microdata for all workers in England and Wales in each year.

Figure 3: Changes in Property Crime Rates and 25th Percentile Wages Across the Police Force Areas of England and Wales, 1975-96



Notes:
1. Averaged (population weighted) annualised changes across 42 police force areas of England and Wales (excluding City of London), 1975-96.
2. Size of text labelling police force areas is proportional to area population.
3. The slope of the regression line fit through the points on the Figure is –1.580 (associated standard error = .756).

Convictions Data

To model P_{at} we use the number of convictions in both the magistrates and crown court for crime against the property, for both men and women, divided by total recorded crime against the property. This is used as the measure of the probability of conviction rather than the official clear up rate, as the number of potential problems with the reported clear up rate are numerous. In the models for the separate components of the property crime rate the probability of conviction for these specific crimes is considered. We also consider some results using average sentence lengths as a deterrence measure.

IV. Empirical Findings

Basic Results

In all models the dependent variable is expressed in log odds form (as suggested by our theoretical specification). The estimation method controls for area fixed effects by including as regressors indicators for each police force area. This is equivalent to using within groups, i.e. all variables are expressed in deviations from their area specific means (taken over time).[18] We also include in all regressions fixed time effects. Hence all impacts are identified as a result of differential changes in the relevant variables across areas and time. Finally in all regressions the observations are weighted by the area and time specific population size. The standard errors reported are robust to heteroskedasticity.

The basic results are presented in Table 1. We present a set of results which we discuss in turn below, where we sequentially add a number of key variables, starting with the column (1) where we include the wage variable (in this case the 25th percentile of the real hourly wage distribution in each police force area). Following this we add the probability of conviction and the share of 15-24 year olds in the population (column (2)). In columns (3) though (5) we allow for the possibility that the deterrence measure is endogenous leading to biased causal inference for both the wage and deterrence impacts. Thus in column (3) we use an alternative measure of deterrence, the average sentence length; in column (4) we use the lagged conviction rate dated (t-2), and in column (5) we use instrumental variables: We instrument the contemporaneous conviction rate using as instruments its (t-2) value as well as the expenditure on police (per police officer).[19] Finally, column (6) reports a dynamic specification including the lagged crime rate. At the bottom of the table we report marginal effects for the log wage measure and the deterrence measure. The marginal effect is the change in the crime rate due to a unit change of the variable in question.

The coefficient reported in column (1) shows a strongly significant impact of the 25th percentile log wage on crime. This clearly points to an important trade off between crime and labour market incentives. The marginal effect reported in the last row of the Table reveals that an area with a 10 percent higher 25th percentile wage is predicted to have a 0.7 percentage point lower crime rate, which viewed against the actual value of the crime rate (about 8 percent, or 80 crimes per 1000 people, in the 1990s) is quite a high number. Including the conviction rate and the population share of young people increases the coefficient on the 25th percentile wage in absolute terms a little (the marginal effect becoming –0.081), such that the negative and significant impact is reinforced. The conviction rate itself has a negative impact as predicted by the model in Section 2 of the paper. The corresponding marginal effect is -.026.

[18] Within groups is consistent (as T becomes large) even in the presence of lagged dependent variables or other such regressors that are not are strictly exogenous. Thus, with a panel of such length (20 years) any bias from using within groups is likely to be minimal (see Nickell, 1981 for example).

[19] Expenditure per police officer is only available from 1980 on. We have experimented with its lagged value as an instrument and we get similar results. However, since this leads to further loss of observations we reported the results based on the longer period obtained with this instrument dated t.

Table 1: Within-Groups Regressions Explaining Property Crime Rates in 42 Police Force Areas, England and Wales, 1975-96

	Property Crime					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	OLS, With Lagged Dep. Var.
Log(25th Percentile Real Hourly Wage)	-1.237 (.189)	-1.506 (.159)	-1.077 (.243)	-1.631 (.192)	-1.640 (.192)	-.324 (.087)
Share of People Aged 15-24 in Population		4.377 (1.221)	4.009 (1.809)	4.388 (1.333)	4.584 (1.347)	1.100 (1.347)
Log(Conviction Rate)		-.486 (.033)		-.816 (.062)	-.814 (.062)	-.127 (.023)
Log(Average Sentence Length)			-.093 (.046)			
Lagged Dependent Variable						.817 (.025)
Police Force Area Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	.939	.961	.936	-	-	.989
Sample size	924	924	750	840	834	924
Marginal effect for the deterrence measure		-.026	-.005	-.045	-.045	-.038
Marginal effect of Log(25th Percentile Real Hourly Wage)	-.066	-.081	-.061	-.090	-.091	-.096

Notes:
1. The dependent variable is Log[PC/(1-PC)] where PC is the property crime rate.
2. Heteroskedastic consistent standard errors in parentheses.
3. All regressions are weighted by area & time population size.
4. The IV results in columns (4) and (5) instrument the log(conviction rate). Instruments used are: column (4) log(conviction rate)(t-2); column (5) log(conviction rate)(t-2), expenditure on police.
5. The marginal effect in column (6) is the implied long run marginal effect.

We also include the share of 15-24 year olds in the area population as an explanatory variable, because most crimes are committed by relatively young people and because this share could be correlated negatively with our wage measure.[20] Including this demographic variable slightly increased the impact of wages on crime, but does not alter the conclusions in any fundamental way.

As already noted one may be concerned about the potential endogeneity of the conviction rate. Column (3) therefore uses an alternative measure of deterrence, the average sentence length. This also has a negative, statistically significant coefficient estimate and again the strong negative 25th percentile wage effect remains. Furthermore, the pattern of results remains intact when one instruments the conviction rate as shown in columns (4) and (5). Because of this we focus on the column (2) (OLS) model as our benchmark model and proceed from there in the empirical investigations that follow. But it should be noted that all results we go on to report go through if we adopt the IV model in our later analyses.

In column (6) we introduce dynamics, to control for peer effects. This reveals a significant persistence in area crime rates, but the wage and conviction effects remain. Furthermore the marginal effect of the wage measure (calculated as the implied long run effect) is very similar to that reported in the earlier columns revealing the very robust nature of the impact of wages on crime. These results imply that because of peer effects (or other sources of persistence) the impact of changes in wages take some time to work through.

[20] Levitt (1999) argues that the characteristics of the whole age distribution should be included for reliable predictions of the impact of demographics on crime. However since it turned out that demographics did not have an important role in identifying incentive effects we have not pursued this here. It should also be noted that the focus of the Levitt paper is rather different to ours. He asks the question of how much changes in age structure can explain changes in aggregate crime rates. On the other hand we are interested in whether area differences in demographics are related to cross area differences in crime rates.

Table 2: Within-Groups Regressions for various crime rates. 42 Police Force Areas, England and Wales, 1975-96

	Vehicle Crime	Theft and Handling	Burglary
	(1)	(2)	(3)
Log(25th Percentile Real Hourly Wage)	-2.533 (.277)	-1.058 (.126)	-1.026 (.147)
Share of People Aged 15-24 in Population	4.876 (2.021)	3.016 (.909)	2.246 (1.132)
Log(Crime Specific Conviction Rate)	-.639 (.044)	-.601 (.026)	-.853 (.040)
Police Force Area Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
R-Squared	.949	.972	.970
Sample Size	924	924	924
Marginal effect for the deterrence measure	-.012	-.023	-.015
Marginal effect of Log(25th Percentile Real Hourly Wage)	-.048	-.040	-.018

Notes:
1. The dependent variable is Log[Crime/(1-Crime)] where Crime is the relevant crime rate.
2. Heteroskedastic consistent standard errors in parentheses.
3. All regressions are weighted by population size.

Table 3: Considering the impact of different wage measures

	Property Crime		
	(1)	(2)	(3)
Log(25th Percentile Real Hourly Wage)	-1.506 (.159)		
Log(10th Percentile Real Hourly Wage)		-1.388 (.222)	
Log(25th Percentile Real Hourly Wage For Retail Trade Workers)			-1.079 (.225)
Share of People Aged 15-24 in Population	4.377 (1.221)	3.261 (1.398)	3.544 (1.425)
Log(Crime Specific Conviction Rate)	-.486 (.033)	-.478 (.035)	-.454 (.037)
Police Force Area Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
R-Squared	.961	.956	.954
Sample Size	924	924	924
Marginal effect for the deterrence measure	-.026	-.026	-.024
Marginal effect of wage measure	-.081	-.074	-.058

Notes:
1. As for column (2) of Table 1.

All in all these results are very much in line with the simple choice theoretic model of Section II. Crime rates are higher in initially high crime areas, but also where wages at the bottom end of the area-specific wage distribution are low and where the probability of being convicted is low.

Different Crime Rates

Table 2 reports estimates of the column (2) Table 1 model for a number of different crime rates. Column (1) considers vehicle crimes whilst columns (2) and (3) separate out the components of property crime, namely theft and handling (column 2) and burglary (column 3). A very similar pattern emerges with the 25^{th} percentile wage always displaying a significant negative association with area crime rates. There are some differences in the magnitudes of the estimated coefficients (and marginal effects), but on the whole the earlier findings are reconfirmed for all crime rates.

Other Wage Measures

So far we have relied on using the 25^{th} percentile wage as our measure of wages at the lower end of the wage distribution that we feel is relevant for people on the margins of crime. There are several reasons why one may wish to also consider other wage measures. First, one might think the 25^{th} percentile is not necessarily the appropriate quantile to consider. Second, one may think of crime as being more attractive for people with less labour market attachment, who may well be working in sectors with low skill jobs with less of a career structure than elsewhere, and hence the wage measure should be better targeted to that labour market.

We consider both of these possibilities in Table 3, which reports results for different wage measures. For ease of comparison the results of the second column of Table 1 are reproduced in column (1) of Table 3. The rest of the Table is then devoted to the alternative wage measures. Column (2) reports results using the 10th percentile wage rather than the 25th percentile. Column (3) uses the 25th percentile wage from the wage distribution of retail trade workers. In all cases the negative impact of wages on crime remains. The 10th percentile wage gives a very similar marginal effect. The marginal effect for the retail wage measure is slightly lower but still exerts a strong statistically significant impact on crime. Moreover, all the marginal effects in the Table are quite large when compared to the crime rate.

Bounding the Wage Effects

The results reported so far may be biased if the variation of wages over time and areas are due to composition changes generated by differential employment or unemployment rates across areas and time, rather than reflecting changes in the price of low skill labour. However, as we noted in Section III we can bound the wage effects by also presenting what one may think of as a 'worst case' scenario where one calculates a wage measure based on the assumption that all those out of work are less productive in the labour market than those in work. Working with the 25th percentile, this adjusted wage measure will be a lower bound to the 25th percentile of the *unconditional* wage distribution (by area and time). In this way, part of the observed wage fluctuation is attributable to compositional changes and is not used to measure changes in incentive effects (over time and across areas).

Table 4 reports results using this 'selection corrected' wage. The resulting marginal effects should be interpreted as the lower bound (in absolute value) of the impact of wages on crime. The results are very reassuring. The coefficient on the corrected wage is negative and always statistically significant. Moreover, the bounds between the uncorrected and this worst case correction seem fairly tight.

Table 4: Using the lower bound on the wage quantile to measure the impact of wages, allowing for composition effects

	Property Crime	
"Selection Corrected Wage Measures"	(1)	(2)
Log(25th Percentile Real Hourly Wage)	-.805 (.121)	
Log(25th Percentile Real Hourly Wage For Retail Trade Workers)		-.701 (.171)
Share of People Aged 15-24 in Population	5.341 (1.527)	4.119 (1.593)
Log(Crime Specific Conviction Rate)	-.466 (.034)	-.445 (.036)
Police Force Area Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
R-Squared	.955	.952
Sample Size	924	924
Marginal effect for the deterrence measure	-.025	-.024
Marginal effect of wage measure	-.043	-.038
Marginal effect of wage measure from "uncorrected models"	-.081	-.058

Notes:
1. As for column (2) of Table 1.

Area Unemployment Rates

In the model of Section II the unemployment rate plays no role as the labour market opportunities available to individuals are summarised by the wage variable. However, many empirical studies of crime do focus upon the crime-unemployment association. In our data associations with unemployment are not very important. There is, in terms of raw data, a positive, statistically significant correlation between area crime rates and the area unemployment rate. But that is as far as it goes. Once one includes area fixed effects into crime equations, there is seen to be no association with area unemployment rates. Our results strongly suggest the relevant area wage measure to be much more important for area crime variations than is the area-specific unemployment rate.[21]

These results are not as surprising as it may first seem, at least according to our model. Crimes are committed by both the employed and the unemployed. An increase in criminal activity may crowd out leisure as well as work.

Including a Measure of r_{at}

The structural model of Section II predicts that a measure of the potential returns to crime, r_{at}, should be included as an independent variable in empirical crime models. This, of course, is a very hard thing to do given the difficulties associated with obtaining measures of the returns to crime. However, we have tried to compute an area measure of crime returns from victimization data reported in the British Crime Surveys (BCS) of 1984, 1988, 1992, 1994 and 1996. We construct our measure from a question (relating to the year before the survey) where crime victims were asked to report the value of stolen property.[22] Of course this measure has some limitations. Nevertheless, we have incorporated a variable measuring the average value of the reported crime into our area crime models for the years of overlap with the BCS value of stolen property data (1983, 1987, 1991, 1993 and 1995).[23]

---

[21] For example, a simple regression of the logistic property crime rate on the unemployment rate alone produced a coefficient (standard error) of 4.425 (.410). But when the unemployment rate is added to the Table 1 column (2) specification it attracted an insignificant coefficient (-0.092, standard error = .369) whilst the 25th percentile wage remained negative and strongly significant (coefficient = -1.593, standard error = 0.150).

[22] The BCS question asked to crime victims is: 'What would you estimate was the total value of what was stolen?'. If respondents asked what this meant, interviewers were told to tell them that to express value as replacement value, and that stolen cheques/credit cards count as no value.

[23] The BCS areas and police force areas do not match perfectly so we have had to match in the BCS data at a higher level of aggregation.

Table 5 reports property and vehicle crime equations including the value of crime measure. Three specifications are reported in each case. The first (columns (1) and (4)) is the basic fixed effects model incorporating only the value of crime measure along with the area and time dummies. For both crime measures there is a positive impact of our returns measure and the crime rates. The bottom of each table reports the marginal effects. From these we see that an area with a 10 percent higher value of crime measure is predicted to have a 0.07 percentage point higher property crime rate and a 0.03 percentage point higher vehicle crime rate.

The models in the remainder of Table 5 additionally include the explanatory variables from the earlier analysis (the 25^{th} percentile wage and the share of young people). For each crime measure, two specifications are presented, one using the wage with no correction for composition effects, the other including the selection corrected wage. The measure of the returns to crime remains positive and significant though inclusion of the other variables does reduce the magnitude of the association with crime. Furthermore, and very reassuringly for our analysis, the wage effects remain much the same as before and are quite tightly bounded by the procedure correcting for composition effects.

Table 5: Including a Measure of the returns to crime

	Property Crimes			Vehicle Crimes		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Value of Crime From BCS Data)	.122 (.049)	.075 (.033)	.085 (.040)	.187 (.074)	.111 (.042)	.136 (.065)
Log(25th Percentile Real Hourly Wage), Uncorrected		-1.166 (.463)			-2.534 (.862)	
Log(25th Percentile Real Hourly Wage), "Selection Corrected"			-.701 (.293)			-.768 (.525)
Share of People Aged 15-24 in Population		2.696 (2.435)	5.107 (2.412)		4.599 (4.408)	9.052 (4.171)
Log(Crime Specific Conviction Rate)		-.596 (.054)	-.581 (.049)		-.802 (.080)	-.748 (.069)
Police Force Area Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	.929	.964	.962	.901	.949	.938
Sample Size	210	210	210	210	210	210
Marginal effect for Log(Value of Crime From BCS Data)	.007	.004	.005	.003	.002	.003
Marginal effect for the deterrence measure		-.032	-.031		-.015	-.014
Marginal effect of wage measure		-.063	-.038		-.047	-.014

Notes:
1. As for Table 2, column (2).
2. The Value of Crime From BCS Data variable is from the 1984, 1988, 1992, 1994 and 1996 British Crime Surveys and is entered into the equations for the previous year. So the equations are estimated for 42 police force areas in 1983, 1987, 1991, 1993 and 1995.

V. Concluding Remarks

Crime against the property has risen sharply in the UK since the late 1970s. This, at least in part, reflects a much longer trend increase in crime (see Home Office, 1998). In this paper we have tried to see whether a simple choice theoretic model of criminal behaviour is in line with the UK experience since the 1970s. The model predicts that crime rates should be higher where and when wages at the bottom end of the wage distribution are lower, reflecting poorer labour market opportunities, where the probability of being caught is lower, where crime rates are already higher and where the potential returns to crime are high. We have presented empirical findings, based on panel data on the police force areas of England and Wales between the mid-1970s and mid-1990s that are in line with the predictions of this model. All variables we tried reflecting incentives for committing crimes have significant and large impacts on crime rates: Increased wages and deterrence measures reduce crimes, while increases in the direct economic returns from crime increase criminal activity.

From a policy perspective the implications of our results is that improvements in human capital accumulation through the education system or other means that can be showed to be effective for enhancing individual labour market productivity, coupled with suitable deterrence measures would be important ingredients in reducing crime. In line with recent research (such as Glaeser, Sacerdote and Scheinkman, 1996) our empirical findings also show how crime rates tend to persist over time across areas where previous crime incidence is high, which can be interpreted as being due to due to peer group or neighbourhood effects.

Appendix

Table A1: Crime Definitions in Notifiable Offences Data

Violence Against the Person	murder, manslaughter, infanticide, attempted murder, threat or conspiracy to murder, child destruction, causing death by dangerous driving or by driving under the influence of drink or drugs, causing death by aggravated vehicle taking, wounding or other endangering of life, endangering a railway passenger, endangering life at sea, other wounding etc., abandoning a child under two years, child abduction, procuring illegal abortion, concealment of birth
Sexual Offences	buggery, indecent assault, indecency between males, rape, unlawful sexual intercourse, incest, procuration, abduction, bigamy, gross indecency with a child
Burglary	burglary in a dwelling or other building, aggravated burglary
Vehicle Crime	theft of motor vehicle, theft from a vehicle, aggravated vehicle taking, vehicle interference and tampering and criminal damage to a vehicle.
Theft and Handling	aggravated vehicle taking, theft from the person, theft in a building, theft by an employee, unauthorised taking from mail, abstracting electricity, theft of a pedal cycle, theft from vehicle, theft from shops, theft from a machine or meter, theft or unauthorised taking of a vehicle, other theft, handling stolen goods
Fraud and Forgery	fraud by company director, false accounting, other fraud, forgery or use of false drug prescription, other forgery

References

Becker, Gary (1968) Crime and punishment: An economic approach, Journal of Political Economy, 76, 175-209.

Blundell, Richard, Amanda Gosling, Hidehiko Ichimura and Costas Meghir (2000) Changes in the distribution of male and female wages accounting for selection and composition effects, Institute for Fiscal Studies mimeo.

Blundell, Richard, Thomas Stoker and Howard Reed (1999) Interpreting movements in aggregate wages: the role of labour market participation, Institute for Fiscal Studies Working Paper 99/13.

Campbell, John Y. and John H. Cochrane (1999) "By force of habit: A consumption based explanation of aggregate stock market behaviour", Journal of Political Economy, 107, 205-51.

Ehrlich, Isaac (1973) Participation in illegitimate activities: A theoretical and empirical investigation, Journal of Political Economy, 81, 521-563.

Fagan, Jeffrey and Richard Freeman (1997) Crime, work and unemployment, mimeo.

Freeman, Richard (1996) Why do so many young American men commit crimes and what might we do about it?, Journal of Economic Perspectives, 10, 125-42.

Freeman, Richard (1999) The economics of crime, in O. Ashenfelter and D. Card (eds) Handbook of Labor Economics, North Holland.

Glaeser, Edward and Bruce Sacerdote (1999) Why is there more crime in cities?, Journal of Political Economy, 107, S225-S258.

Glaeser, Edward, Bruce Sacerdote and Jose Scheinkman (1996) Crime and social interactions, Quarterly Journal of Economics, 111, 507-48.

Gosling, Amanda, Stephen Machin and Costas Meghir (1999) The changing distribution of male wages, forthcoming Review of Economic Studies.

Gould, Eric, Bruce Weinberg and David Mustard (1998) Crime rates and local labour market opportunities in the United States: 1979-95, mimeo.

Grogger, Jeffrey (1998) Market wages and youth crime, Journal of Labor Economics, 16, 756-91.

Hansen, Kirstine (1999) The role of education, labour market and social factors in shaping the age crime profile, mimeo.

Heckman, James and Guilherme Sedlacek (1985) Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market, Journal of Political Economy, 93, 1077-1125.

Home Office (1998) Reducing offending: as assessment of research evidence on ways of dealing with offending behaviour, Home Office Research Study 187.

Johnson, George and Frank Stafford (1999) International trade and the labor market, in O. Ashenfelter and D. Card (eds) Handbook of Labor Economics, North Holland.

Katz, Lawrence and David Autor (1999) Wage inequality, in O. Ashenfelter and D. Card (eds) Handbook of Labor Economics, North Holland.

Patrick Langan and David Farrington (1998) Crime and justice in the United States and England and Wales 1981-96, US Bureau of Justice Statistics.

Levitt, Steven (1998) Juvenile Crime and Punishment, Journal of Political Economy, 106, 1156-85.

Levitt, Steven (1999) The limited role of changing age structure in explaining aggregate crime rates, Criminology, 37, 581-97.

Machin, Stephen (1996) Wage Inequality in the UK, Oxford Review of Economic Policy, 12:2: 47-62.

Machin, Stephen (1998) Recent shifts in wage inequality and the wage returns to education in Britain, National Institute Economic Review, 166, 87-96.

Manski, C (1994) "The Selection Problem" in Sims C (Eds) Advances in Econometrics; Sixth World Congress Vol. 1 pp 143 - 170.

May, Sean (1999) Wages and youth arrests, MIT mimeo.

Meghir, Costas and Edward Whitehouse (1996) The evolution of wages in the UK: evidence from micro data, Journal of Labor Economics.

Nickell, Stephen (1981) Biases in dynamic models with fixed effects, Econometrica, 49, 1417-26.

Reilly, Barry and Robert Witt (1996) Crime, deterrence and unemployment in England and Wales: An empirical analysis, Bulletin of Economic Research, 48, 137-59.

Witt, Robert, Alan Clarke and Nigel Fielding (1998) Crime, earnings inequality and unemployment in England and Wales, Applied Economics Letters, 5, 265-67.