USING PANEL DATA TO CHARACTERISE PRISONER AND DEMOGRAPHIC AGE CHARACTERISTICS

By Wee Lee Loh[1]

ABSTRACT

Using panel data to study the macro-linkage between demographic and prisoner age characteristics this empirical paper investigates the relationship between age and the probability of being sentenced via an alternative framework. Fixed Effects GLS was used on both an unbalanced panel and a balanced subpanel data set. Both level and log transformed models were tested. Indeed, there is evidence that younger people are more than proportionately committed to the prisons. This result is only significant when the unbalanced panel is used. This exercise also illustrates the possible dangers of creating a balanced subpanel from an unbalanced data set.

---

1 The author is currently a candidate for a Masters in Arts in Statistics in the Department of Statistics at Harvard University. The author likes to thank his thesis supervisor, Tim Vogelsang for much guidance and feedback. All of the errors in this paper are solely due to the author. Financial support from the College of Arts and Sciences Undergraduate Research Fund and Cornell Economics Department is gratefully acknowledged. For comments, please email: wlloh@fas.harvard.edu or via phone at 607-261-6234.

1. Introduction

After Nobel Laureate Gary Becker first published his pioneering article "Crime and Punishment: An Economics Approach", many economists have forayed into the field of criminology. The primary motivation for the study of crime is obvious. The expected social cost of criminal activities is always negative. By learning how the frequency and intensity of crime can be reduced, the expected social loss can be better minimised, which in turn enhances aggregate social welfare.

Historically, the causes and origins of crime have been the subjects of investigation by many disciplines, including economics. While no definitive conclusions have yet been reached, a number of factors affecting the volume and type of crime that occurs from place to place have been delineated. These factors include economic conditions, cultural conditions and effective strength of the judicial system. Amongst the spectrum of accepted explanations, this paper focuses on one particular factor, the demography.

The relationship between demography and criminology dates back to the eighteenth century, when social statisticians, particularly Quetelet and Lexis (Stigler, 1986) drew heavily on both criminological and demographic data. Between the 1960s and 1970s, criminologists noticed the value in relating demographic patterns to crime rates. Chilton & Spielberger (1971), Ferdinand (1970) and Wellford (1973) are some of the many authors[2] who studied this relationship. In these studies, crime rates were regressed onto demographic variables and test if the crime rate was significantly

[2] More authors include Sagi & Wellford (1968), Lee (1984), Cohen & Land (1987). This list is by no means exhaustive.

explained by those demographic variables. Instead, the proposed model in this paper will instead regress the median age of prisoner onto the median age of the population.

There is little debate over the basic form of the relationship between age and crime. Age-specific crime rates tend to rise with increasing age during childhood and early adolescence, peak in the early 20s and then fall rapidly thereafter (Messner & Rosenfeld, 1999). This form of relationship is derived from the lifecourse trajectory literature, which authors like Nagin and Land (1993), Nagin, Farrington and Moffitt (1995) have extensively researched. This paper will not analyse this crime-age relationship via a life-course perspective. Although the general patterns pertaining to individual age and crime are well established, controversies about the specific nature of the patterns and the interpretations exist.

Hirschi and Gottfredson (1983) first touched off considerable debate several years ago when they published an article asserting that the relationship between age and crime is invariant with respect to social and economic circumstances. They claim that the aging of the criminal is the sole explanation for the decline in criminal offending. On the other hand, Greenberg (1977, 1983) offered an alternative structural interpretation of the relationship between crime and age. His single explanatory framework encompasses elements of both strain and control theories of crime. Unlike Gottfredson and Hirschi who view criminal behaviour as rooted in the biological process of aging, Greenberg (1983) posits that the age-crime relationship is highly contingent on the changing social position of adolescents.

Against this backdrop of age-crime relationships, some of the linkages between crime and demography will be discussed. Messner and South (2000) explore many ways in which criminal and demographic behaviours are reciprocal, at both the microsocial and macrosocial levels, even though there is no single comprehensive theory that associates demographic trends and criminal behaviour. They assert that age is one of the most powerful and robust individual-level factors for predicting criminal behaviour. This finding constitutes the primary motivation to focus on the age-crime relationship so as to better understand the root of criminality.

Researchers have proffered some possible age-demographic linkages. One possible link is through compositional change. Evidence[3] consistently indicates that younger people are at comparatively higher risk of becoming offenders. Aggregate crime rates vary as a function of the changes in the age composition of the population. If the high offending age classes expand, then the total crime rates will increase correspondingly, even if no change occurs in age-specific rates (Messner & South, 2000). Besides compositional links, contextual (or causal) associations exist. Contextual associations emphasize the effects of varying age group sizes on crime rates. Economist Richard Easterlin, prominently associated with this work, hypothesizes that changes in birth cohort sizes will alter the economic and social context within which people grow up. In particular, he focuses on the consequences of the large baby boom cohort on both the labour market and social institutions, which jointly affect the strain and control influences discussed by Greenberg. Like the compositional explanation, the contextual effect is consistent with Greenberg's social structural explanation.

Given that individual (micro) level predictors have been well-documented, this paper aims to study if the same level of predictability exists at the macrosocial level.

[3] Messner & Rosenfeld (1999), Sampson & Lauritsen, (1994)

Previous research have shown that, on average, one particular person is more likely to commit a crime in his youth than when the same person is older, holding external circumstances invariant. Instead of focusing on a particular youth, this paper desires to generalize this age-crime relationship to a collective group of individuals; the study of a collective group behaviour amounts to what is termed as macro-social. Messner and Sampson (1991) discovered that the indicators of age composition do not always relate to crime rates in the predicted manner, suggesting that the generalising of individual characteristics to the macro level is far more complicated than expected. The contribution of this paper will be to propose an alternative model to explicate the relationship between the age of prisoners and the age of the population. This model will then be used to determine if younger people are, on average, more probable to commit crime than older people.

Aside from understanding the crime-age relationship in itself, this study can have useful implications in charting crime-fighting public policy. Previous studies have concluded that age is one of the strongest individual predictors for crime offending (Messner & Rosenfeld, 1999). Will this micro-level relationship remain valid at the macro-level? For example, if a country's population is aging, then will it necessarily face a declining crime rate? If so, a government can strategically reduce its budgetary allocation (assuming constant returns to scale, i.e. effectiveness of crime control per unit dollar) to crime fighting departments, like the police department and judiciary systems, in view of an expected decline in crime. Researchers have effectively used changes in population composition, especially age composition to offer forecast about future crime trends in crime based on demographic patterns. This was successfully done to explicate the crime trends in the US from the 1960s to 1980s. Instead of focusing on one particular country, a cross-country analysis is performed in this study, to investigate if the same age-crime relationship exists in different countries. With data from 32 countries, a panel data set is used to study the age-crime relationship. This is important because previous studies have not devoted sufficient attention to such cross-country analysis, but instead, focused on the US. This paper hopes to bridge this gap in the existing literature.

The results of this paper confirm the hypothesis that younger people have a higher probability of entering the prisons system. This relationship is only significant when the unbalanced data set is used. This finding is robust to the model specifications as both level and log forms generated significant results. However, the thesis is not supported when the balanced subpanel is substituted for the unbalanced panel. When a balanced subpanel is created from an unbalanced panel, important information may be cast out.

This paper will be organized as follows. Section 2 elucidates the theoretical considerations made in constructing the model while Section 3 contains the summary statistics derived from the data set. Next, section 4 contains the results and discussion of this paper. Finally, the conclusion and appendix will be included at the end.

2. Theoretical Considerations

Certain terms will be used consistently throughout this paper. Criminals refers to all persons associated with the commission of a behaviour legally defined as a crime (Feldman, 1977). This avoids issues with the definition of criminality, even though criminologists proffer different views on it. To avoid any confusion between the statistical notion of 'population' and the demographic use of 'population', the usage of 'population' will be restricted to the demographic context. Prisoners are used as a convenient proxy for criminals. A working definition for a younger person is one who has age less than the median age of a prison cohort in a given year.

The aim of this paper is to revisit the hypothesis that younger people are at higher risk of entering the penitentiary system. Instead of using a model with crime rate as the dependent variable, this paper will propose an alternative framework for analyzing the same issue. Bramer and Piquero (2002) argue that the usage of crime rate as a dependent variable is not exactly the most suitable for a crime-age study. A disproportionate amount of crime is generated by a relatively small cohort of chronic offenders, which skews the distribution of crime.

The following paragraphs will motivate the set-up of the model and illustrate how the population interacts with the prison cohort. In a starting year t, there is a given cohort of prisoners. The members of this group can be systematically arranged by their individual ages to generate an age distribution F_t. Using this age distribution Ft, a median age value of M_t can be obtained. From the year t to year $t + 1$, the new age distribution F_{t+1} is determined by two principal factors, the inflow and outflow of prisoners in year t. If there are no changes in the type and size of the existing prison cohort from year t, then the median age in year $t+1$, M_{t+1} will simply be $M_t + 1$. To characterise the age distribution, median statistic is used instead of the mean statistic, because the median is less sensitive to the effects of outliers.

Each year, there will be some prisoners who leave; this group of departing prisoners will be termed as the outflows. The outflow of prisoners comprises two groups. The first group consists of prisoners who die naturally in year t while the second group includes those who have just completed their sentences. Just as there are departing prisoners, there will also be fresh inmates entering the prisons; these newly arriving prisoners will be labelled as the inflows. The rate of this inflow will directly alter the size of the prison cohort in year *t+1*, while the age of the inflowing prisoners will change the age-distribution of prisoners F_{t+1}.Suppose an inflowing prisoner has an age less than that of M_t, then his age will decrease the value of M_{t+1} such that $M_{t+1} < M_t + 1$, ceteris paribus.

The ages of both the incoming and out-flowing prisoners will directly affect the new age distribution F_{t+1} in year $t + 1$. Suppose in a year t, the net inflow of prisoners is 10. Five of these inflows have ages above M_t while the other five have ages less than M_t. As expected, the prisoner at the median quantile remains the same, and the new median age in year $t + 1$ will still be $M_t + 1$. Instead, if 2 of them have ages above M_t while the other eight have ages less than M_t, it is obvious that the new median age will be less than or equal to $M_t + 1$.

Besides the age characteristic, the distribution F_{t+1} is also affected by the rate in which prisoners flow in and out of the system. The larger the net exchange of prisoners, the larger the effect will be on F_{t+1}. The rate of prisoner inflow is largely determined by the changes in incarceration rate in year *t*. The incarceration rate in year t is the total number of prisoners over the total population size. A change in incarceration rate is used as a proxy for changes in the net inflow of prisoners.

The net inflow of prisoners is the difference between the rate of inflowing prisoners and the rate of out-flowing prisoners. Since this exchange takes place between the prisons and the population, the age variation of the population will directly affect the age distribution of the prison cohort. In each year, there is an imaginary "selection vector" that picks members from the population to be part of the inflowing prisoners. The goal of this paper is to determine if this selection vector systematically chooses the younger people from the population to enter into penal institutions. If this is true, it will reflect the hypothesis that younger people are more probable to be convicted and sentenced to the prisons. Although the selection vector is not directly observed, the inference is done on the changes in M_t from year to year. If Mt decreases yearly, it will be reasonable to deduce that the selection vector has been picking proportionately greater number of young people to enter the prisons.

2.1 The Model

Abstracting from this theoretical framework, the following corresponding regression model is proposed. The model is given by equation (1).

$$\text{prisonage}_{it} = \beta_0 + \beta_1\,\text{popage}_{it} + \beta_2\,\text{prisonsize}_{it} + \beta_3\,\text{popsize}_{it} + \beta_4\,\text{incarceration}_{it} + \beta_5\,\text{year93}_t + \beta_6\,\text{year94}_t + \beta_7\,\text{year95}_t + \beta_8\text{year96}_t + \beta_9\,\text{year97}_t + \beta_{10}\,\text{year98}_t + a_i + \varepsilon_{it} \quad (1)$$

where t = years 1992 to 1998, i = 1 to N where N denotes the number of countries included in the sample. $Prisonage_{it}$ is defined as the median age of prisoners while $popage_{it}$ is similarly defined as the median age of the population. $Prisonsize_{it}$ denotes the cohort size of prisoners while $popsize_{it}$ denotes the population size. The variable a_i is the unobserved idiosyncratic country effect. This country effect is assumed to encompass the social structural conditions existing in those countries and to be fixed over time. The variables {*year93, year94, year95, year96, year97, year98*} are known as the year effects. These year effect variables account for any systematic events that occur in all of the countries in a given year.

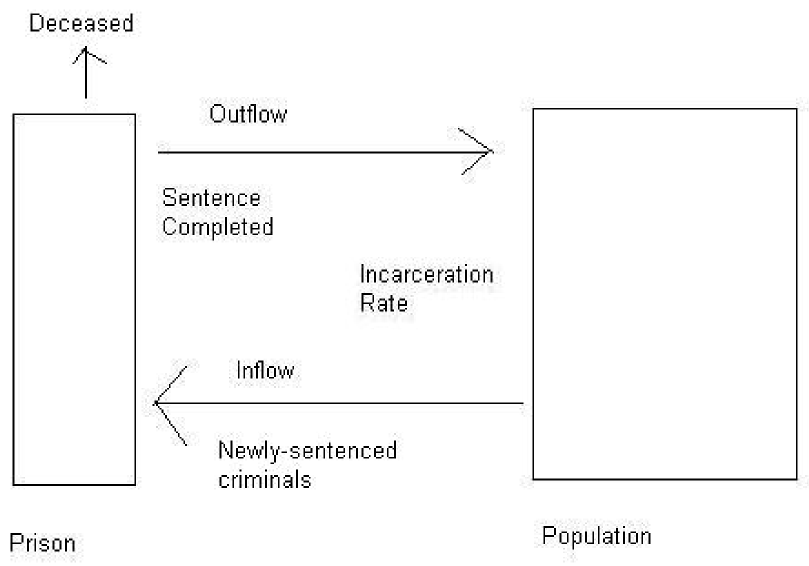


Figure 1: Diagram of Prisoner-Population Interaction

As a robustness check on the modelling specification, the same analysis will be repeated with a log-transformed model. This method is intuitively appealing because all of the data values are positive. The model is then recast into equation (2).

$$\log(prisonage_{it}) = \delta_0 + \delta_1 \log(popage_{it}) + \delta_2 \log(prisonsize_{it}) + \delta_3 \log(popsize_{it}) + \delta_4 \log(incarceration_{it}) + \delta_5\, year93_t + \delta_6\, year94_t + \delta_7\, year95_t + \delta_8\, year96_t + \delta_9\, year97_t + \delta_{10}\, year98_t + a_i + \varepsilon_{it} \quad (2)$$

Besides introducing an alternative modelling specification, the value of this paper also lies in the cross-country nature of investigation. Previous studies[4] frequently used US-based data, which are conveniently available and adequately complete[5]. Many of the countries in the data set are European, which have different legal systems from that of the

[4]See Cohen & Land (1987), Fox (1978), Steffensmeier & Harer (1991, 1999)

[5]See Cohen & Land (1987)

US. It would be a fruitful exercise to see if the same age-crime relationship is detectable in these countries. The list of countries included in this study is shown in table 1.

Table 1: List of Countries

Albania	Lithuania
Austria	Luxembourg
Belgium (B)	Macedonia
Bulgaria	Moldova
Canada	Netherlands (B)
Cyprus (B)	Norway
Czech Republic (B)	Poland
Estonia	Portugal (B)
Finland (B)	Romania
France (B)	Russia
Greece	Slovak (B)
Hungary	Slovenia
Iceland	Spain (B)
Ireland	Sweden (B)
Italy (B)	Switzerland (B)
Latvia	Turkey

Note: (B) denotes the countries included in the balanced subpanel. For elaboration on the balanced subpanel, refer to section 2.2

2.2 Data

The data was retrieved from two main sources. “Penological Information Bulletin No. 18 - 22, Council of Europe” supplied the data on the following variables: prisoner’s median age, the total number of prisoners and the incarceration rate in each country. As the data from this source was limited, it restricted both the scope of countries and the time periods of investigation. The Internet site of the US Census Bureau, Population Division, International Programs Center6 (Table 094 Midyear population, by Age and Sex) provided the data for the other two variables, population median age and population

[6] http://www.census.gov/ipc/www/idbsprd.html

size. The median age of the population was not readily available; manual calculations were done to extract the median figures from the histogram population data.

The entire data set is an unbalanced and incomplete panel. Originally, the data comprised 39 countries, which totalled 213 observations. An observation is considered incomplete if a value is missing for one or more of the variables. To clean up the data set, all incomplete observations were purged. In so doing, the data set was reduced by 66 observations to obtain a complete set of unbalanced panel data which features 32 countries. An unbalanced panel is one in which the individual time series for a different countries have different lengths. The unbalanced panel data set includes all complete data points, where number of countries (N) = 32, time periods (T) = 7, NT (sample size) = 147. In the original set of 39 countries, 15 of the countries had less than 6 years of data.

The results of this study will be valid under the assumption that the missing data for some country *i* is not correlated with the idiosyncratic errors, ε_{it} is applicable. In other words, this assumption rules out any possibility a country will systematically withhold prisoner's age information. However, this assumption is not verifiable by the given data, which is one of the limitations of this study. At best, the justification for this assumption can only be an intuitive argument. If a country had intentionally withheld data on prisoner's age in a certain year, then it will be equally likely to withhold the same data in all other years. If this assumption is false, the inference can be misleading because it is no longer representative of the underlying statistical population.

A balanced subpanel was extracted from the unbalanced data set. In the unbalanced data set, some countries only provided 2 to 3 years of data. Only countries which have at least 6 out of 7 years of data are included in the balanced set (Number of

Countries = 12, Time Periods = 7, NT (sample size = 77). The balanced subpanel constitutes 52% of the unbalanced data set. For details of the specific countries included in the balanced subpanel, please refer to table 1. Table 2 provides a histogram on the number of countries which have specific years of observations. This overview will give a sense of how unbalanced the data set is.

Table 2: Histogram of the number of countries versus the length of the time series

Length of time series	Unbalanced Panel	Balanced Subpanel
1	1	-
2	4	-
3	4	-
4	6	-
5	5	-
6	7	7
7	5	5

A panel data structure is useful because it permits the unobserved effects a_i to be correlated with the explanatory variables. However rich the model is, it will not be able to capture all of a country's characteristics. Criminologists like Worrall and Pratt (2004) advocate the use of panel analysis to control for unobserved individual population heterogeneity. Failure to account for unobserved heterogeneity would lead to bias estimates when cross-jurisdictional studies are done (Cherry, 1999).

A panel data structure permits reasonably good inferences to be drawn on the results. Intuitively, if a relationship between age-crime is universally true, then it will remain valid in different circumstances. Panel study accounts for such varied circumstances. Besides facing diverse socioeconomic conditions, each country has a unique judicial system. If, amidst these differences, the data reveals a relationship between the variables of interest, then there is indeed evidence to establish the investigated relationship. Some of the other advantages of a panel approach as discussed in Baltagi (2001) can be summarized as follows: (1) Panel data give more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency. (2) Panel data are better able to identify and measure effects that are simply not detectable in pure cross-section or pure time-series data (3) Panel data allows us to construct and test more complicated behavioural models. (4) Panel data are able to study the dynamics of adjustment.

2.3 Assumptions

In order to draw valid conclusions from the model, some assumptions need to be made. The precision of the interpretation of the regression results is contingent on the legitimacy of the following assumptions.

Assumption 1: Stability of the relationships over time, i.e. the β_j's are time invariant. This particular model specification implicitly assumes that the underlying relationship between the independent and dependent variables is stable over time, i.e. β_js are not time varying.

One of the valid concerns in any time-based model is that the coefficients estimates are indeed time trending. A simple cursory check for this is to observe the within R2 values. As the empirical results reveal very low within R2 values, it is reasonable to conclude that the coefficients of the variables do not vary over time. In addition to this simple inspection, a more systematic Chow test was implemented to verify the admissibility of this assumption. A Chow (1986) test can be performed to detect if indeed, the coefficient estimates are stable over time. Under Chow's proposed test, the time dummy variables will be interacted with the explanatory variables. For the unbalanced panel, there is sufficient degree of freedom to interact all of the independent variables at once.

Assumption 2: Sample observations are geographically independent: $E(a_i a_j) = 0$ *for all* $i \neq j$. This assumption describes how the unobserved heterogeneity of country i, a_i is not correlated with the unobserved effects of country j, a_j . As described earlier, the term a_i captures the structural similarities that exist within a country over the period of analysis, 1992 to 1998. If this assumption is false, biasness will be introduced into the model.

As noted above, this paper uses both an unbalanced panel and a balanced subpanel. The data is considered unbalanced or incomplete when the same country is not observed over the entire time period. The concern with the incompleteness arises out of worries if the missing data is systematically linked with the variables specified in the model. If so, there will be biasness in the estimates of the coefficients of the explanatory variables. In the econometric literature, this issue is commonly known as selection bias in panel data.

For a random draw i from the underlying statistical population, let $s_i \equiv (s_{i1}, s_{i2}, \ldots, s_{iT})'$ denote the $T \times 1$ vector of selection indicators: $s_{it} = 1$ if (x_{it}, y_{it}) is observed and zero otherwise. Let

$$\mathbf{w}_{it} = x_{it} - T_i^{-1} \sum\nolimits_{r=1}^{T} s_{ir}\, x_{ir} \qquad \text{and} \qquad T_i = \sum\nolimits_{t=1}^{T} s_{it}$$

Assumption 3:

(a) $E(\varepsilon_{it} | x_i, s_i, c_i) = 0$. This assumption posits that the error term in model equation (1) is not systematically linked to the vector of selection indicators.

(b) E(sitex0itexit) is non-singular. This assumption permits the matrix containing all explanatory variables and selection vector to be invertible, much like the full-rank assumption of a typical least squares model.

If all of these assumptions are legitimate, then the estimates of the fixed effects model on the unbalanced panel will be consistent. Furthermore, if the assumption (a) and (b) are valid given this data set, then the results for the random effects will also be reliable. Wooldridge (chapter 10, 2002) contains some details on these Fixed Effects assumptions.

Fixed effects GLS regression was estimated for on both the unbalanced panel and balanced subpanel data sets. This method require a slightly different set of assumptions, they both permit an arbitrary correlation to exist between the unobserved country heterogeneity, a_i and the explanatory variables in any time period.

This unobserved heterogeneity includes other demographic features like race, education level or socioeconomic conditions. A brief illustration of how each of the explanatory variable in the model is possibly correlated with the unobserved effect, a_i, will follow. If indeed such correlation exists, then it will justify the use of a fixed effects method to obtain the coefficient estimates.

The variables, $popage_{it}$ and $popsize_{it}$ are correlated with the birth rate of a country, which is part of a_i. When the birth rate of a population is decreasing over time coupled with a decreasing death rate, the proportion of older people increases. This will naturally raise the median age of the population. Similarly, it is obvious that as the birth rate increases, the corresponding population size must expand, ceteris paribus.

The second variable, $prisonsize_{it}$ is directly related to the crime rate of a country. The higher the crime rate is, the larger the number of criminals is. With more criminals, given a constant probability of sentencing convicted criminals to prison facilities, there will be a larger inflow into the prisons. The current stock of prisoners is largely determined by the cumulative historic crime rate in the country.

The incarceration rate is also affected by the inherent preferences of the judicial system. Some judicial systems lean towards deterrent sentencing, to award tougher punishments on criminals while other systems are more lenient towards younger offenders, believing that the young person still has a redeemable future.

If all of the above-mentioned unobserved factors do not vary over time, then my estimates will be more meaningful and reasonably more accurate. Many of these country-specific characteristics like birth rate or attitude towards crime take a long time to change, especially so given the short time frame of analysis (1992 to 1998).

3. Summary Statistics

Table 3 provides an overview of the median ages across all the countries in each time period. By juxtaposing the two data sets, the quantitative differences between them can be easily observed. The weighted population median age is calculated by weighting the median age of a particular country with its population size. Similarly computed, the weighted prison median age is derived by weighting the median age of a country's prison cohort with the size of its prison cohort.

Table 3 –Weighted Median Age of the Prisoners and Population

	Weighted Median Age							
	Unbalanced Data			Balanced Data			Difference	
Year	Population		Prisoners	Population		Prisoners	Population	Prisoners
1992	35.3	>	31.1	35.6	>	31.1	-0.3	0
1993	34.8	>	31.0	35.2	>	30.3	-0.4	0.7
1994	35.4	>	32.1	36.2	>	31.7	-0.8	0.4
1995	36.2	>	31.8	36.6	>	31.8	-0.4	0
1996	35.9	>	32.2	37.0	>	32.5	-1.1	-0.3
1997	34.3	=	34.3	37.0	>	32.5	-2.7	1.8
1998	34.5	<	36.2	37.4	>	32.4	-2.9	3.8

Note: Difference = Unbalanced figure – Balanced figures

Although the data sets exhibit a consistently non-decreasing trend over time for both populations and prisoners figures, the regression results do not provide sufficient evidence to indicate any trending behaviour. This is a real concern because trending variables can produce spurious regressions. *Stylized Observation: Younger people are more than proportionally represented in the prisons.*

This observation concurs with previous evidence which indicates that young people are at comparatively high risk of becoming offenders (Messner & Rosenfeld, 1999, Sampson & Lauritsen, 1994), considering that judicial systems only subject persons above ages of 18 or 21 to penitentiary systems. When the weighted median average age of prisoners is 32.7 (unbalanced data), it shows that persons between 18 (or 21 depending) to 32.7 constitute 50% of the prison population. On the other hand, 50% of the population is aged between 0 and 36.5. By comparing these two ranges, it is acceptable to conclude that younger people are more than proportionally represented in the prisons systems. The median age of prisoners is observed to be lower than that of the population for most of the time periods. The exceptions are for years 1997 and 1998. In these two years, the unbalanced data set includes Turkey, which is featured only in 1997 and 1998. Turkey has an unusually high median age of prisoners, 49 (1997) and 51(1998), which significantly raised the weighted median ages for those 2 years. For a summary statistics of each variable, please refer to tables A1 and A2 in the Appendix.

The main difference between this stylized observation and the thesis of this paper is the typical stock versus flow differentiation. This observation pertains to the stock of prisoners up to the year 1992 while the thesis examines the flow of prisoners entering the penitentiary from 1992 to 1998.

To facilitate a graphical overview of the data points, there are two scatter plots of $prisonage_{it}$ and $popage_{it}$ (refer to Appendix Graphs 1 and 2). In graph 1 (Unbalanced data set), there is a dense cluster of points in the top left corner of the box. This cluster hints of an association between prison and population age characteristics. When $prisonage_{it}$ lies in the range of 30 to 35 years, it is reasonable to expect to lie between 30 and 40 years. Two of the outliers in the bottom right corner belong to Turkey, which features twice in the data set. Unlike Graph 1, Graph 2, which plots the balanced data, reveals a possible linear association, instead of a dense cluster of points. This linearity is observed in the area where $popage_{it}$ ranges from 35 to 40 and $prisonage_{it}$ ranges from 30 – 35. This illustration partially explains why subsequently only the models ran using the balanced data set consistently reveals a significant relationship between $prisonage_{it}$ and $popage_{it}$ whereas those regression exercises that used the unbalanced data set did not.

4. Empirical Analysis

Fixed effects GLS regressions were estimated for both equations (1) and (2). The following table lists the coefficients estimates of the independent variables. The entire regression exercise was repeated on the balanced subpanel. For most of the coefficients, the estimates of the balanced subpanel are of the same order of magnitude as its counterpart in the unbalanced data set.

Table 4 – Coefficient Estimates of Equation (1) and (2)

Dependent variable: $prisonage_{it}$

	Unbalanced		Balanced	
	Level	Log	Level	Log
Constant	14.93	30.71**	-0.337	39.41**
	(19.90)	(10.64)	(30.21)	(18.20)
$popage_{it}$	-0.05	0.0147	0.104	0.551
	(0.197)	(0.162)	(0.767)	(0.820)
$prisonsize_{it}$	7.19e-05	0.00639	6.54e-05	-0.022
	(1.24e-04)	(0.0554)	(1.38e-04)	(0863)
$popsize_{it}$	1.23e-06	-1.75**	1.6e-06	-2.35**
	(1.29e-06)	(0.686)	(1.44e-0.6)	(1.15)
$incarceration_{it}$	-0.0159	-0.079	-0.0201	0.0045
	(0.0163)	(0.0533)	(0.0226)	(0.0736)
$year93_t$	0.0445	0.0153	-0.382	-7.14e-04
	(0.667)	(0.0204)	(0.814)	(0.0260)
$year94_t$	-0.286	0.0148	-0.305	9.78e-03
	(0.695)	(0.0220)	(0.729)	(0.0243)
$year95_t$	0.639	0.0455**	-0.288	0.0199
	(0.730)	(0.0227)	(0.864)	(0.0294)
$year96_t$	0.573	0.0519**	0.216	0.0439
	(0.695)	(0.0232)	(1.005)	(0.0348)
$year97_t$	0.629	0.0565**	0.923	0.0714*
	(0.742)	(0.0241)	(1.14)	(0.0399)
$year98_t$	1.02	0.0724**	-0.198	0.0396
	(0.811)	(0.0261)	(1.24)	(0.0429)
between R^2	0.0671	0.0842	0.010	0.0120
within R^2	0.0843	0.1241	0.1561	0.1793
Overall R^2	0.0493	0.0237	0.0028	0.0039

Notes: Numbers in parentheses are std errors. All estimates are rounded to 2 or 3 significant figures.
** Significant at 5% for coefficient estimate of the variable = 0.
*Significant at 10% for coefficient estimate of the variable= 0

Across all the estimation methods, no explanatory variable is always statistically significant. In other words, there is no relationship between two variables that is robust across both data sets and both model specifications. In two cases, $popsize_{it}$ has a statistically and practically significant coefficient estimate under the log specifications. It shows that if the population size expands by 1%, the corresponding prisoner's median age will decrease by 1.75% (balanced) or 2.75% (unbalanced). One possible explanation for this negative correlation is that with expanding populations, there will naturally be a greater number of younger people. As younger people have higher propensity to commit crimes, there will be more of these young people in the prisons, which will then lower the observed median ages of the prison cohorts.

Based on the motivation of this study, it is of interest to test if the coefficient estimate of $popage_{it}$ is less than 1. The interpretation of this test is that the aging process in the prisons is slower than the aging process of the population. In other words, the net inflows entering the prisons are much younger than the prison cohort's median age. Thus, the hypothesis test is set up as follows:

$H_0: \beta_1 \geq 1$ $\qquad$ $H_1 : \beta_1 < 1$

Table 4: Results of Hypothesis Test

	Unbalanced level	Unbalanced log	Balanced level	Balanced log
Coefficient	-0.05	0.0147	0.104	0.551
T-stat	-5.33	-6.08	-1.168	-0.548
P-value*	4.92E-08	5.953E-10	0.121	0.292

* = taken under Gaussian distribution

For the above regressions, there was sufficient evidence to reject the null H_0 when the unbalanced data was used. However, under both specifications, the balanced subpanel did not provide any significant result. This observation is consistent with Baltagi's (2001) concern. He cautions that by creating a balanced subpanel nested within the unbalanced panel is only throwing away important information. In this case, these results confirm Baltagi's postulation that the relationship that is detectable in the unbalanced panel might disappear in the balanced subpanel.

Both models using the unbalanced data produced significant estimates in the same order of magnitude $O(10^{-2})$ although both coefficient estimates have different signs. To a certain extent, the difference in signs can be attributed to the noise in the data, especially since the difference is merely 0.06. More importantly, the interpretation of this estimate is that indeed, the median age of a prison cohort is rather invariant to changes in demographic age distribution. This is contrary to what is typically expected. It means that even as the general population exhibit an aging phenomenon, aging is absent in the prison cohort. This then provides direct evidence that younger people are more than proportionately sent to the prisons.

4.1 Chow's Test for Assumption 1

Under the Chow Test, equation (3) is estimated using fixed effects GLS to validate assumption 1. To analyse if a parameter β is time invariant is to separately estimate 7 different cross-sections for each year and to test if the coefficient estimate is the same throughout all the years. Since a panel data structure is available, it would be better to work within the panel framework than to revert to cross-sectional analysis. The following equation is proposed to test the equivalent hypothesis. This method interacts the variable of interest with time dummies, to see if a different coefficient estimate is produced in each time period. To write the following equation in a compact fashion, I will define the following terms:

$$interactpopage_{it} = \alpha_1 \, year93_t \, popage_{it} + \alpha_2 \, year94_t \, popage_{it} + \alpha_3 \, year95_t \, popage_{it} + \alpha_4 \, year96_t \, popage_{it} + \alpha_5 \, year97_t \, popage_{it} + \alpha_6 \, year98_t \, popage_{it}$$

$$interactprisonsize_{it} = \pi_1 \, year93_t \, prisonsize_{it} + \pi_2 \, year94_t \, prisonsize_{it} + \pi_3 \, year95_t \, prisonsize_{it} + \pi_4 \, year96_t \, prisonsize_{it} + \pi_5 \, year97_t \, prisonsize_{it} + \pi_6 \, year98_t \, prisonsize_{it}$$

$$interactpopsize_{it} = \rho_1 \, year93_t \, popsize_{it} + \rho_2 \, year94_t \, popsize_{it} + \rho_3 \, year95_t \, popsize_{it} + \rho_4 \, year96_t \, popsize_{it} + \rho_5 \, year97_t \, popsize_{it} + \rho_6 \, year98_t \, popsize_{it}$$

$$interactincarceration_{it} = \tau_1 \, year93_t \, incarceration_{it} + \tau_2 \, year94_t \, incarceration_{it} + \tau_3 \, year95_t \, incarceration_{it} + \tau_4 \, year96_t \, incarceration_{it} + \tau_5 \, year97_t \, incarceration_{it} + \tau_6 \, year98_t \, incarceration_{it}$$

$$prisonage_{it} = \eta_0 + \eta_1 \, popage_{it} + \eta_2 \, prisonsize_{it} + \eta_3 \, popsize_{it} + \eta_4 \, incarceration_{it} + \eta_5 \, year93_t + \eta_6 \, year94_t + \eta_7 \, year95_t + \eta_8 \, year96_t + \eta_9 \, year97_t + \eta_{10} \, year98_t + \eta_{11} \, interactpopage_{it} + \eta_{12} \, interactprisonsize_{it} + \eta_{13} \, interactpopsize_{it} + \eta_{14} \, interactincarceration_{it} + c_i + \xi_{it} \quad (3)$$

Joint Significance Test

H_0: $\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 0$

H_1: H_0 is not true

The intuition for this joint significance test is that, if all of the variability in $prisonage_{it}$ has already been captured by the year effects and the variable $popage_{it}$, then the time-interacted variables will contain no explanatory power. This is equivalent to have zero coefficient estimates. A simple Likelihood Ratio test is implemented for this exercise. This test is subsequently repeated for all other explanatory variables by modifying equation (3). The hypothesis result from all tests is shown in table 5.

Table 5: Hypothesis Test of Parameter Stability (Chow Test)

Data	Test Stat	Dist	P-value
Unbalanced	42.15	Chi-square (24)	0.0124
Balanced	46.16	Chi-square (24)	0.0042

In both cases, there is violation of assumption 1 that the value of β is not stable over time. In other words, if the analysis is restricted to individual cross-sections, there will be a different set of β for each time period. However, it is reasonable to argue that this changing β is not driving the results obtained in table 4. If the changing β did affect the hypothesis testing, then it would affect both the balanced and unbalanced data sets identically. Since this Chow test encompasses the interaction of all explanatory variables with time dummies, it is possible that the source of changing β is observed in other variables apart from $popage_{it}$.

5. Conclusions

In this paper, a fixed effects GLS regression was estimated to study the age-crime relationship, to see if younger people have a higher probability of being sentenced to the prisons. Under a larger unbalanced panel, the data presents sufficient evidence to support the thesis that younger people are more likely to be sentenced to the prisons. But this result disappears when the balanced subpanel data is used instead. Baltagi (2001) argues that whenever a balanced subpanel is created from an unbalanced panel, important information is thrown away. In this empirical exercise, this argument is validated because the significant relationship observed under the unbalanced panel becomes insignificant under the balanced subpanel.

This paper advances the study of age-crime relationship via an alternative model. The use of a different approach permits a robustness check that a covariance indeed exists between criminal's age and demographic age. This is evidence that previous micro level observations of an age-crime linkage remain true at the macro level. Further research can be done using non-parametric methods to estimate the model. Strong structural assumptions were imposed by the use of a typical linear panel model. The model can also be enriched by including more country-specific characteristics, such as the type of judicial system, size of police forces, or arrest rates or factors describing socioeconomic conditions. Similarly, other demographic factors like birth rate, death rate will be useful inclusions as well.

REFERENCES

Baltagi, Badi H (2001). Econometric Analysis of Panel Data, 2nd Ed. Chichester: John Wiley & Sons.

Cherry, T. L. (1999). Unobserved heterogeneity bias when estimating economic model of crime. *Applied Economic Letters* 6: 753 – 757.

Chilton, R., & Spielberger, A. (1971). Is delinquency increasing? Age structure and the crime rate. *Social Forces, 49*, 487-493.

Chow, Gregory (1986). Random and Changing coefficient models, In Z. Griliches and M.D Intrilligator (eds.), vol II, Handbook of Econometrics. North-Holland, Amsterdam, 1213-1245.

Cohen, LE & Land, KC (1987). Age structure and crime: symmetry versus asymmetry and the projection of crime rates through the 1990s. *American Social. Review 52,* 170-183.

Feldman, M. Philip (1977). *Criminal Behavior: A Psychological Analysis*. London: John Wiley & Sons.

Ferdinand, T. N. (1970). Demographic shifts and criminality: An inquiry. *British Journal of Criminology, 10*, 169-175.

Fox, JA (1978). *Forecasting Crime Data: An econometric analysis*. Lexington, MA: Lexington Books.

Greenberg, David F. (1977). Delinquency and the Age Structure of Society. In *Contemporary Crises 1*:189-223.

Greenberg, David F. (1983). "Crime and Age" Pp 30-35. In Levinson, David (Ed), *Encyclopedia of Crime and Justice*, vol 1. New York: Macmillan.

Hirschi, Travis and Gottfredson, Michael (1983). *Causes of Delinquency*. Berkeley, CA: University of California Press.

Lee, G. W. (1984). Are crime rates increasing? A study of the impact of demographic shifts on crime rates in Canada. *Canadian Journal of Criminology, 26*, 29-41.

Messner Steven F. & Rosenfeld R. (1999). Social structure and homicide: theory and research. In Smith, MD & Zahn, MA (eds), *Homicide: A Sourcebook of Social Research*, pp. 27–41. Thousand Oaks, CA: Sage.

Messner, Steven F. & Sampson, RJ (1991). The sex ratio, family disruption and rates of violent crime: the paradox of demographic structure. Soc. Forc. 69, 693-713.

Nagin, D. S., Farrington, D. P., & Moffitt, T. E. (1995). Life-course trajectories of different types of offenders. *Criminology, 33*, 111-140.

Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed poisson model. *Criminology, 32*, 327-362.

Sagi, P. C., & Wellford, C. F. (1968). Age composition and patterns of change in criminal statistics. *Journal of Criminal Law, Criminology & Police Science, 59*, 29-36.

Short, JF Jr. 1971. Introduction. In Short, JF Jr (ed), *The Social Fabric of the Metropolis: Contributions of the Chicago School to Urban Sociology*, pp. xi-xlvi. Chicago: University Chicago Press.

Sampson RJ & Lauritsen JL. (1994). Violent victimization and offending: individual-, situational-, and community-level risk factors. In Reiss, AJ Jr & Roth, JA (eds), *Understanding and Preventing Violence 3*:1 – 114. Washington, DC: Natl. Acad. Press

South, Scott & Messner, Steven, F. (2000). Crime and Demography: Linkages through age. Multiple linkages, reciprocal relations. *Annual Review of Sociology* 26, 83-106.

Stark, R. (1981). Deviant places: a theory of the ecology of crime. *Criminology 25,* 893-909.

Steffensmeier, D. Allen, E. A,Harer, M. D & Streifel, C. (1989). Age and the distribution of crime. *American Journal of Sociology 94*, 803-831.

Steffensmeier, D. & Harer, MD (1991). Did crime rise of fall during the Regan presidency? The effects of an "aging" US population on the nation's crime rate. *J. Res. Crime Delinquency 28*, 330-59.

Steffensmeier, D. & Harer, MD (1999). Making sense of recent US crime trends, 1980 to 1996/1998: age composition effects and other explanations. *J. Res. Crime Delinquency* 36:235-74.

Wellford, C. F. (1973). Age composition and the increase in recorded crime. *Criminology, 11*, 61-70.

Wooldrige, Jeffrey (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press

Worrall, John. & Pratt, Travis C. (2004). On the consequences of ignoring unobserved heterogeneity when estimating macro-level models of crime. *Social Science Research 33*, 79-105

Appendix A

Table A1 – Summary of variables (Unbalanced data set)

Variable	Min	Max	Average	Weighted Average
prisonage	24	54	31.8	32.7
popage	21.1	39.3	34.8	35.2
prisonsize	101	664700	19530.4	19530.4
popsize	259012	148409753	15352462	15352462
incarceration	24.7	443	109.6	198.1

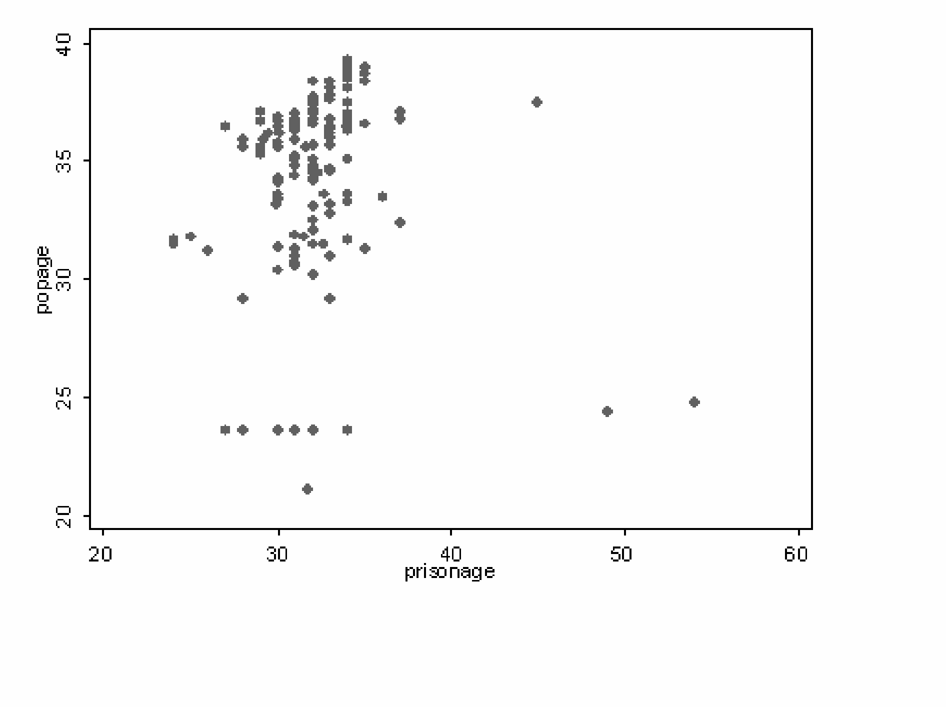
Note: Prisoner variable is weighted according to prison size. Population variables are weighted according to population size.

Table A2 – Summary of variables (Balanced data set)

Variable	Min	Max	Average	Weighted Average
prisonage	25	37	32.0	31.8
popage	23.6	39.3	34.8	36.5
prisonsize	157	54442	17476.9	17476.9
popsize	707825	58866290	18234392	18234392
incarceration	24.7	214	90.3	71.3

Note: Prisoner variable is weighted according to prison size. Population variables are weighted according to population size.

Graph 1 - Scatter Plots of $prisonage_{it}$ **versus** $popage_{it}$ **(unbalanced data set)**



Graph 2 Scatter Plots of $prisonage_{it}$ **versus** $popage_{it}$ **(balanced data set)**